



1/78

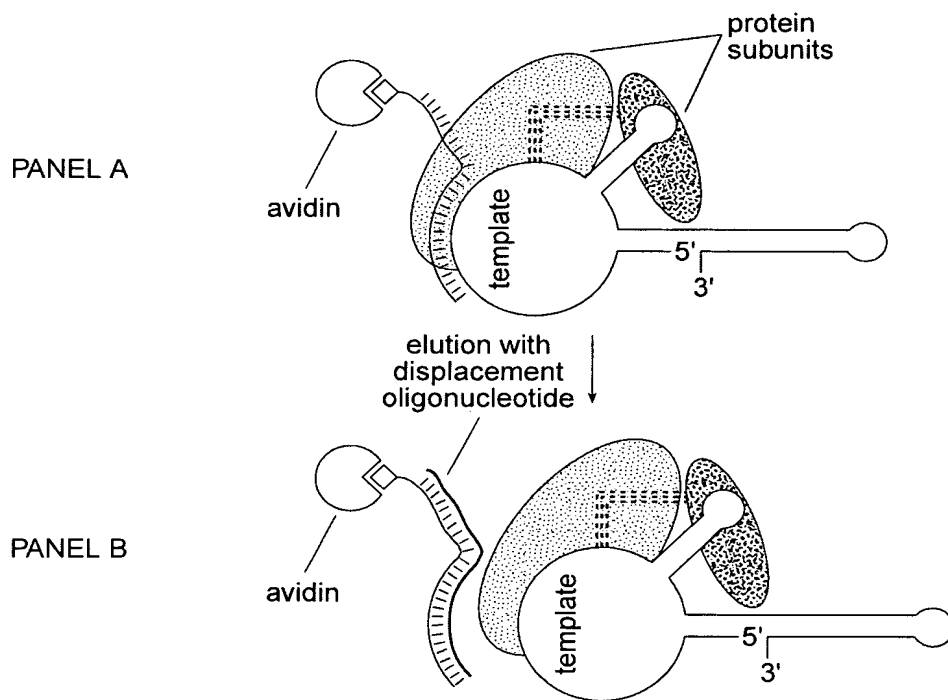


FIG. 1

+



2/78

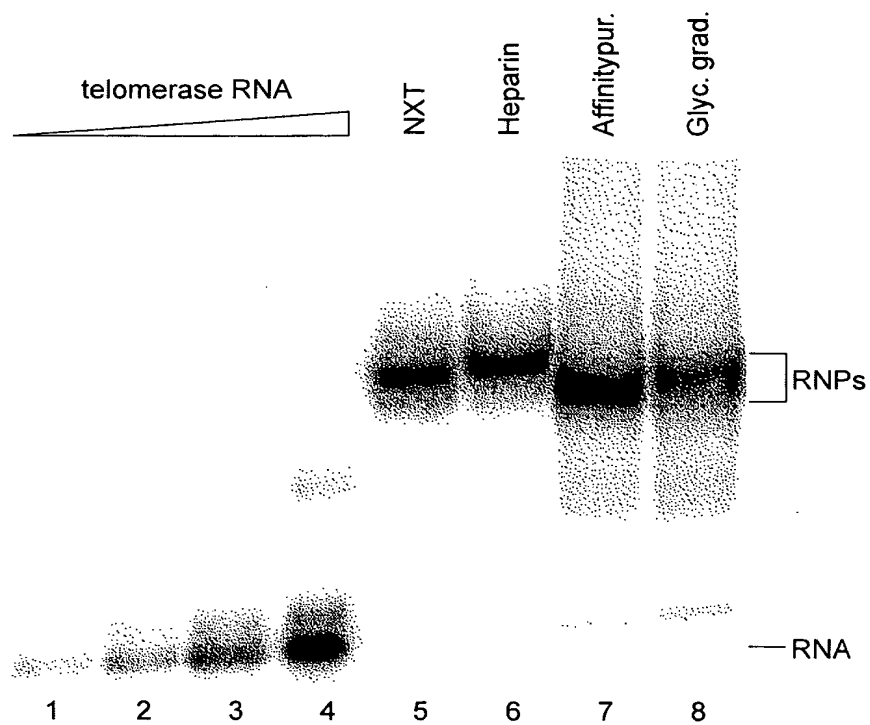


FIG. 2

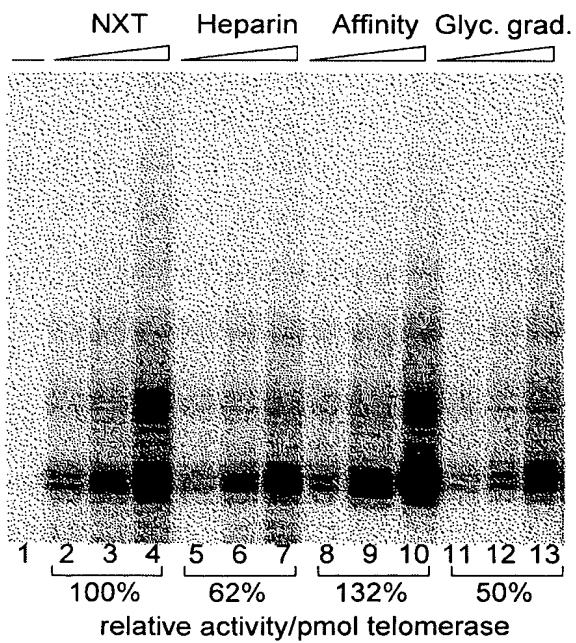
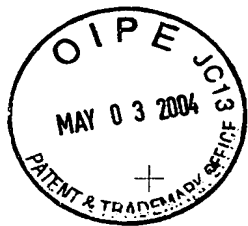


FIG. 3

+



3/78

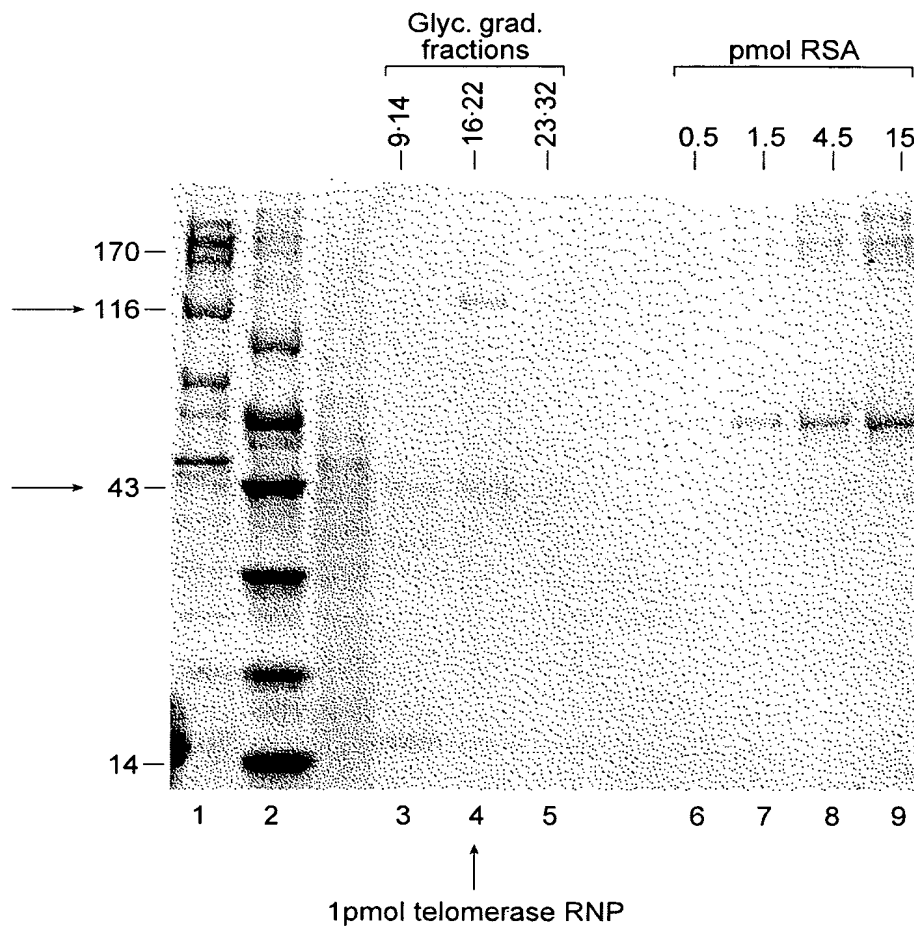


FIG. 4

+



4/78

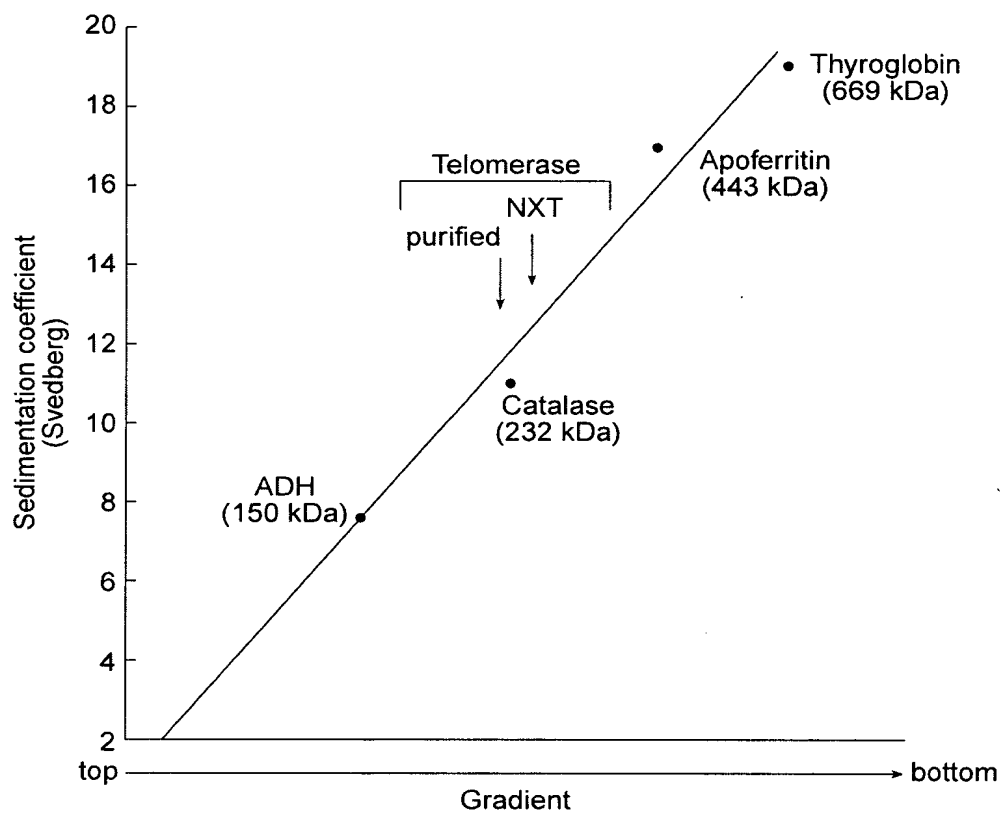


FIG. 5



5/78

Telomerase:

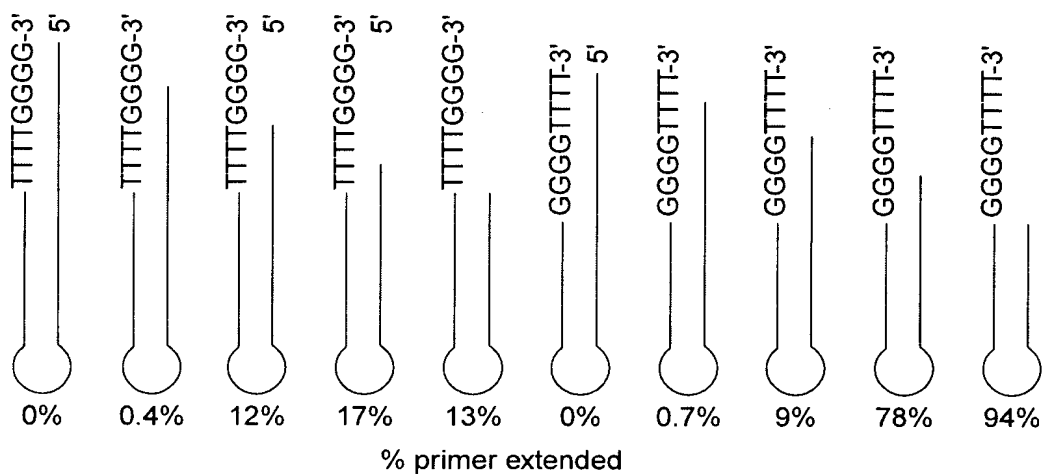
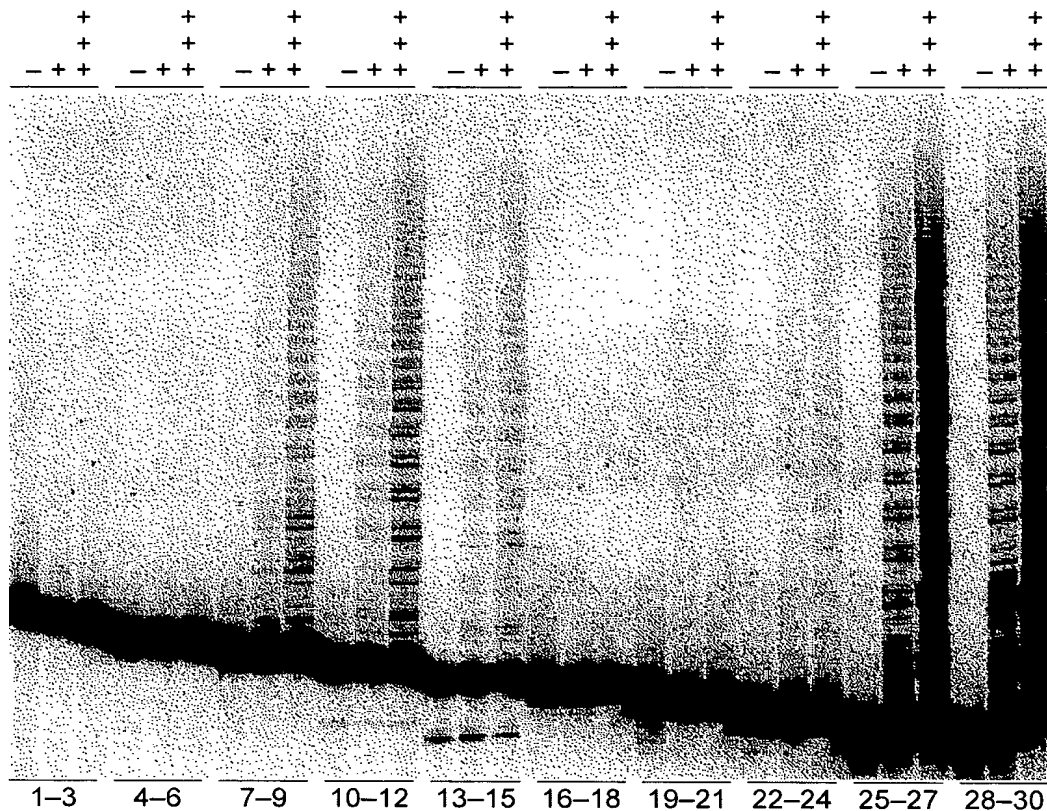


FIG. 6

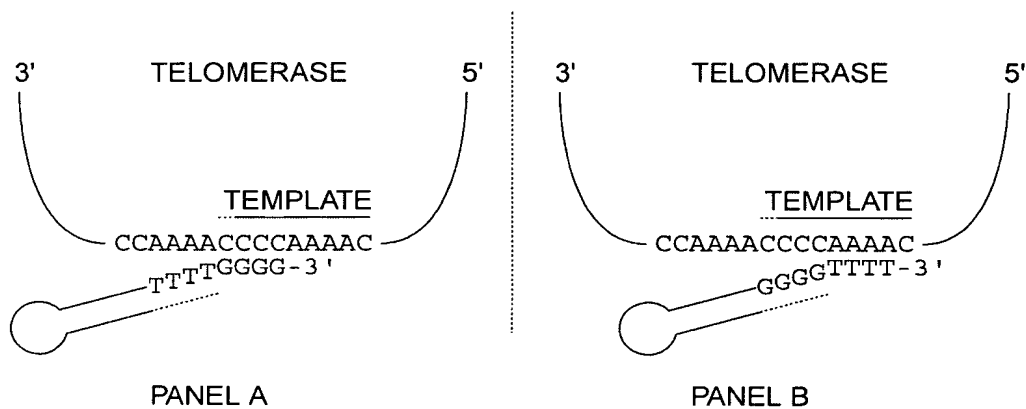


FIG. 7

1	CCCCAAAACC	CCAAAACCCC	AAAACCCCTA	TAAAAAAAGA	AAAAATTGAG
51	GTAGTTTAGA	AATAAAATAT	TATTCCTGCA	CAAATGGAGA	TGGATATTGA
101	TTTGGATGAT	ATAGAAAAT	TACTTCCTAA	TACATTCAAC	AAGTATAGCA
151	GCTCTTGTAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTCGAAA
201	TCGCCTTCAT	TGACTATTCC	AAAGTTGCAA	AAACAATTAG	AGTTCTACTT
251	CTCGGATGCA	AATCTTTATA	ACGATTCTTT	CTTGAGAAAA	TTAGTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAAATTGAAA	CATTACTAAT	GTTTAAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTTT	AGATCACTTC	TTAAGGAGCA
401	TTATGGAGAA	AATTACTTAA	TACTAAAAGG	TAAACAGTTT	GGATTATTTT
451	CCTAGCCAAC	AATGATGAGT	ATATTAAATT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACTT	ACCAAAGACA	AACTCGCTAT	AAAACGCAAG
551	AAAAAGTTTG	ATAATCGAAC	AGCAGAAGAA	CTTATTGCAT	TTACTATTCT
601	TATGGGTTTT	ATTACAATTG	TTTTAGGTAT	CGACGGTGAA	CTCCCGAGTC
651	TTGAGACAAT	TGAAAAAGCT	GTTTACAAC	GAAGGAATCG	CAGTTCTGAA
701	AGTTCTGATG	TGTATGCCAT	TATTTTGTGA	ATTAATCTCA	AATATCTTAT
751	CTCAATTTAA	TGGATAGCTA	TAGAAAACAA	CCAAATAAAC	CATGCAAGTT
801	TAATGGAATA	TACGTAAAT	CCTTTGGGAC	AAATGCACAC	TGAATTTATA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAT	GCTTTAGAGA	CTGATTTAGC
901	TTACAACAGA	TTACCTGTTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGCGAAATGA	AAAGAAGACT	AAAGAAAGAG	ATTTCAAAT
1001	TTGTTGATTC	TTCTGTAACC	GGAATTAACA	ACAAGAATAT	TAGCAACGAA
1051	AAAGAAGAAG	AGCTATCACA	ATCCTGATTC	TTAAAGATTT	CAAAAATTCC
1101	AGGTAAGAGA	GATACATTCA	TTAAAATTCA	TATATTATAG	TTTTTCATTT
1151	CACAGCTGTT	ATTTTCTTTT	ATCTTAACAA	TATTTTTTGA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAATAAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCAT	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACA
1301	GCAGTCATCC	GTTTAAAAAA	TAGTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCCGAAATCT	TAATCAAAAA	GAATTGCGTC	GATATTGCAA
1401	AAGAATCGAA	CTCTAAATCT	TTCGTTAATA	AGTATTACCA	ATCTTGATTG
1451	ATTGAAGAGA	TTGACGAGGC	AACTGCACAG	AAGATCATT	AAGAAATAAA
1501	GTAACTTTTA	TTAATTAGAG	AATAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAAA
1601	ATACAAACCT	TGGTCAAAAT	ATTGAGGAAG	GAAAAGAAGA	CCAGTTAGCA
1651	AAAGAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAAGT	GAAGAAATAA
1701	AAGATTTAT	TTTTTCAATA	ATTTATTGAA	AAGAGGGGTT	TTGGGGTTTT
1751	GGGGTTTTTG	GG			

FIG. 11



7/78



FIG. 8

+



1 AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
251 TTAGAAGATA TTAAATATTT TGCGCAGACA AATATTGTTG CTACTCCACG
301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTC
451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTCAA
501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
601 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT
901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
1001 CTTGAGAAAG TCAAAGATTT TAACCTCAAC TACTATTTAA CAAAATCTTG
1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAACCTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
1151 TACACAACCTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
1251 AAAAGAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTCAC
1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACCTGA
1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451 CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAA TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAGA CCGGAAGACT ACAAATTAAT CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTTGGAATC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTTGTGTTG CAAATGGAAG CAAGTTGGAC AACCAAAAC CTCTTTTGCA
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
1951 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTT TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
2051 TTTAGAAAAG AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101 TGCACCTGAA GGAGGACAAT ATCCAACCTT ATTCAAGTGT CTGAAAAATG
2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTATT TTAAGAAAAG TAACCTACTT CAACCAGTCA TTAATATTTG
2251 CCAATATAAT TACATTAAC TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT
2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 9A



9/78

2401	CCCTGAAAAT	CCAAATGTTA	ATCTTCTAAT	GAGACTTACA	GATGACTATC
2451	TTTTGATTAC	AACTCAAGAG	AATAATGCAG	TATTGTTTAT	TGAGAAACTT
2501	ATAAACGTAA	GTCGTGAAAA	TGGATTTAAA	TTCAATATGA	AGAAACTACA
2551	GACTAGTTTT	CCATTAAGTC	CAAGCAAATT	TGCAAAATAC	GGAATGGATA
2601	GTGTTGAGGA	GCAAAATATT	GTTCAAGATT	ACTGCGATTG	GATTGGCATC
2651	TCAATTGATA	TGAAAACCTCT	TGCTTTAATG	CCAAATATTA	ACTTGAGAAT
2701	AGAAGGAATT	CTGTGTACAC	TCAATCTAAA	CATGCAAACA	AAGAAAGCAT
2751	CAATGTGGCT	CAAGAAGAAA	CTAAAGTCGT	TTTTAATGAA	TAACATTACC
2801	CATTATTTTA	GAAAGACGAT	TACAACCGAA	GACTTTGCGA	ATAAACTCT
2851	CAACAAGTTA	TTTATATCAG	GCGGTTACAA	ATACATGCAA	TGAGCCAAAG
2901	AATACAAGGA	CCACTTTAAG	AAGAACTTAG	CTATGAGCAG	TATGATCGAC
2951	TTAGAGGTAT	CTAAAATTAT	ATACTCTGTA	ACCAGAGCAT	TCTTTAAATA
3001	CCTTGTTGTC	AATATTAAGG	ATACAAATTT	TGGAGAGGAG	CATTATCCAG
3051	ACTTTTTCCT	TAGCACACTG	AAGCACTTTA	TTGAAATATT	CAGCACAAAA
3101	AAGTACATTT	TCAACAGAGT	TTGCATGATC	CTCAAGGCAA	AAGAAGCAAA
3151	GCTAAAAAGT	GACCAATGTC	AATCTCTAAT	TCAATATGAT	GCATAGTCGA
3201	CTATTCTAAC	TTATTTTGGA	AAGTTAATTT	TCAATTTTGT	TCTTATATAC
3251	TGGGGTTTTG	GGGTTTTGGG	GTTTTGGGG		

FIG. 9B

1	MEVDVDNQAD	NHGIHSALKT	CEEIKEAKTL	YSWQKQVIRC	RNQSQSHYKD
51	LEDIKIFAQT	NIVATPRDYN	EEDFKVIARK	EVFSTGLMIE	LIDKCLVELL
101	SSSDVSDRQK	LQCFGFQLKG	NQLAKTHLLT	ALSTQKQYFF	QDEWNQVRAM
151	IGNELFRHLY	TKYLIFQRTS	EGTLVQFCGN	NVFDHLKVND	KFDKKQKGGG
201	ADMNEPRCCS	TCKYNVKNEK	DHFLNNINVP	NWNNMKSRT	IFYCTHFNRN
251	NQFFKKHEFV	SNKNNISAMD	RAQTIFTNIF	RFNRIRKKLK	DKVIEKIAM
301	LEKVKDFNFN	YYLTKSCPLP	ENWRERKQKI	ENLINKTREE	KSKYYEELFS
351	YTDDNKCVTQ	FINEFFYNIL	PKDFLTGRNR	KNFQKKVKKY	VELNKHELH
401	KNLLLEKINT	REISWMQVET	SAKHFYFDH	ENIYVLWKL	RWIFEDLVVS
451	LIRCFYVTE	QOKSYSKTY	YRKNIDVIM	KMSIADLKKE	TLAEVQEKEV
501	EEWKKS LGFA	PGKLRLIPKK	TTFRPIMTFN	KKIVNSDRKT	TKLTTNTKLL
551	NSHMLKTLK	NRMFKDPFGF	AVFNYDDVMK	KYEEFVCKWK	QVGQPKLFFA
601	TMDIEKCYDS	VNREKLSTFL	KTTKLLSSDF	WIMTAQILKR	KNNIVIDSKN
651	FRKKEMKDYF	RQKFQKIALE	GGQYPTLFSV	LENEQNDLNA	KKTLIVEAKQ
701	RNYFKKDNLL	QPVINICQYN	YINFNGKFYK	QTKGIPQGLC	VSSILSSFFY
751	ATLEESSLGF	LRDESMNPEN	PNVNLLMRLT	DDYLLITTQE	NNAVLFIKEL
801	INVSRENGFK	FNMKKLQTSF	PLSPSKFAKY	GMDSVEEQNI	VQDYCDWIGI
851	SIDMKTLALM	PNINLRIEGI	LCTLNLMQ	KKASMWLKKK	LKSFLMNNIT
901	HYFRKTITTE	DFANKTLNKL	FISGGYKYM	CAKEYKDHFK	KNLAMSSMID
951	LEVSKI IYSV	TRAFFKYLVC	NIKDTIFGEE	HYPDFFLSTL	KHFIEIFSTK
1001	KYIFNRVCM	LKAKEAKLKS	DQCQSLIQYD	A	

FIG. 10



10/78

CCCCAAAACCCCAAAACCCCAAAACCCCTATAAAAAAGAAAAATTGAGGTAGTTTAGA
1 -----+-----+-----+-----+-----+-----+-----+ 60
GGGGTTTTGGGGTTTTGGGGTTTTGGGGATATTTTTTTCTTTTTTAACTCCATCAAATCT

a P Q N P K T P K P L * K K K K L R * F R -
b P K T P K P Q N P Y K K R K N C G S L E -
c P K P Q N P K T P I K K E K I E V V * K -

AATAAAATATTATTCCTGCACAAATGGAGATGGATATTGATTTGGATGATATAGAAAATT
61 -----+-----+-----+-----+-----+-----+ 120
TTATTTTATAATAAGGGCGTGTTCCTCTACCTATAACTAAACCTACTATATCTTTTAA

a N K I L F P H K W R W I L I W M I * K I -
b I K Y Y S R T N G D G Y C F G C Y R K F -
c * N I I P A Q M E M D I D L D D I E N L -

TACTTCCTAATACATTCAACAAGTATAGCAGCTCTTGAGTGACAAGAAAGGATGCAAAA
121 -----+-----+-----+-----+-----+-----+ 180
ATGAAGGATTATGTAAGTTGTTTCATATCGTCGAGAACATCACTGTTCTTTCCTACGTTTT

a Y F L I H S T S I A A L V V T R K D A K -
b T S * Y I Q Q V * Q L L * * Q E R M Q N -
c L P N T F N K Y S S S C S D K K G C K T -

CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCCAAAGTTGCAAAAACAATTAG
181 -----+-----+-----+-----+-----+-----+ 240
GTAACCTTAGACCGAGCTTTAGCGGAAGTAAGTAAAGGTTTCAACGTTTTTGTTAATC

a H C N L A R N R L H C L F Q S C K N N * -
b I E I W L E I A F I D Y S K V A K T I R -
c L K S G S K S P S L T I P K L Q K Q L E -

AGTTCTACTTCTCGGATGCAAACTTTTATAACGATTCTTTCTTGAGAAAATTAGTTTTAA
241 -----+-----+-----+-----+-----+-----+ 300
TCAAGATGAAGAGCCTACGTTTAGAAATATTGCTAAGAAAGAACTCTTTTAATCAAAATT

a S S T S R M Q I F I T I L S C E N * F * -
b V L L L G C K S L * R F F L E K I S F K -
c F Y F S D A N L Y N D S F L R K L V L K -

AAAGCGGAGAGCAAAGAGTAGAAATTGAAACATTACTAATGTTTAAATAAAATCAGGTAA
301 -----+-----+-----+-----+-----+-----+ 360
TTTCGCCTCTCGTTTCTCATCTTTAACTTTGTAATGATTACAAATTTATTTTAGTCCATT

a K A E S K E * K L K H Y * C L N K I R * -
b K R R A K S R N C N I T N V * I K S G N -
c S G E Q R V E I E T L L M F K * N Q V M -

TGAGGATTATTCTATTTTTTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA
361 -----+-----+-----+-----+-----+-----+ 420
ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCTCGTAATACCTCTTTTAATGAATT

a C G L F Y F L D H F L R S I M E K I T * -
b E D Y S I F * I T S * G A L W R K L L N -
c R I I L F F R S L L K E H Y G E N Y L I -

FIG. 12A



11/78

421 TACTAAAAGGTAAACAGTTTGGATTATTTCCCTAGCCAACAATGATGAGTATATTTAAATT
-----+-----+-----+-----+-----+ 480
ATGATTTTCCATTTGTCAAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTTAA

a Y * K V N S L D Y F P S Q Q C C V Y * I -
b T K R * T V W I I S L A N N D E Y I K F -
c L K G K Q F G L F P * P T M M S I L N S -

481 CATATGAGAATGAGTCAAAGGATCTCGATACATCAGACTTACCAAAGACAAACTCGCTAT
-----+-----+-----+-----+-----+ 540
GTATACTCTTACTCAGTTTCCTAGAGCTATGTAGTCTGAATGGTTTCTGTTTGAGCGATA

a H M R M S Q R I S I H Q T Y Q R Q T R Y -
b I C E C V K G S R Y I R L T K D K L A I -
c Y E N E S K D L D T S D L P K T N S L * -

541 AAAACGCAAGAAAAAGTTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTCG
-----+-----+-----+-----+-----+ 600
TTTTGCGTTCTTTTCAAACCTATTAGCTTGTCGTCTTCTGAATAACGTAAATGATAAGC

a K T Q E K V C * S N S R R T Y C I Y Y S -
b K R K K K F D N R T A E E L I A F T I R -
c N A R K S L I I E Q Q K N L L H L L F V -

601 TATGGGTTTTATTACAATTGTTTTAGGTATCGACGGTGAACCTCCCGAGTCTTGAGACAAT
-----+-----+-----+-----+-----+ 660
ATACCCAAAATAATGTTAACAAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA

a Y G F Y Y N C F R Y R R C T P E S C D N -
b M G F I T I V L G I D G E L P S L E T I -
c W V L L Q L F * V S T V N S R V L R Q L -

661 TGAAAAAGCTGTTTACAACCTGAAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT
-----+-----+-----+-----+-----+ 720
ACTTTTTTCGACAAATGTTGACTTCCTTAGCGTCAAGACTTTCAAGACTACACATACGGTA

a C K S C L Q L K E S Q F C K F * C V C H -
b E K A V Y N * R N R S S E S S D V Y A I -
c K K L F T T E G I A V L K V L M C M P L -

721 TATTTTGTGAATTAATCTCAAATATCTTATCTCAATTTAATGGATAGCTATAGAAACAA
-----+-----+-----+-----+-----+ 780
ATAAAACACTTAATTAGAGTTTATAGAATAGAGTTAAATTACCTATCGATATCTTTGTTT

a Y F V N * S Q I S Y L N L M D S Y R N K -
b I L * I N L K Y L I S I * W I A I E T N -
c F C E L I S N I L S Q F N G * L * K Q T -

781 CCAAATAAACCATGCAAGTTTAAATGGAATATACGTTAAATCCTTTGGGACAAATGCACAC
-----+-----+-----+-----+-----+ 840
GGTTTATTTGGTACGTTCAAATTACCTTATATGCAATTTAGGAAACCCTGTTTACGTGTG

a P N K P C K F N G I Y V K S F G T N A H -
b Q I N H A S L M E Y T L N P L G Q M H T -
c K * T M Q V * W N I R * I L W D K C T L -

841 TGAATTTATATTGGATTCTTAAAGCATAGATACACAGAATGCTTTAGAGACTGATTTAGC
-----+-----+-----+-----+-----+ 900
ACTTAAATATAACCTAAGAATTTTCGTATCTATGTGTCTTACGAAATCTCTGACTAAATCG

a C I Y I G F L K H R Y T E C F R D C F S -
b E F I L D S * S I D T Q N A L E T D L A -
c N L Y W I L K A * I H R M L * R L I * L -

FIG. 12B



12/78

TTACAACAGATTACCTGTTTTGATTACTCTTGCTCATCTCTTATATCTTTAAAAGAAGCA
901 -----+-----+-----+-----+-----+ 960
AATGTTGTCTAATGGACAAAACCTAATGAGAACGAGTAGAGAATATAGAAATTTTCTTCGT

a L Q Q I T C F D Y S C S S L I S L K E A -
b Y N R L P V L I T L A H L L Y L * K K Q -
c T T D Y L F * L L L L I S Y I F K R S R -

GGCGAAATGAAAAGAAGACTAAAGAAAAGAGATTTCAAAATTTGTTGATTCTTCTGTAACC
961 -----+-----+-----+-----+-----+ 1020
CCGCCTTACTTTTCTTCTGATTTCTTTCTCTAAAGTTTAAACAACCTAAGAAGACATTGG

a G E M K R R L K K E I S K F V D S S V T -
b A K * K E D * R K R F Q N L L I L L * P -
c R N E K K T K E R D F K I C * F F C N R -

GGAATTAACAACAAGAATATTAGCAACGAAAAAGAAGAGCTATCACAATCCTGATTC
1021 -----+-----+-----+-----+-----+ 1080
CCTTAATTGTTGTTCTTATAATCGTTGCTTTTCTTCTTCTCGATAGTGTTAGGACTAAG

a G I N N K N I S N E K E E E L S Q S * F -
b E L T T R I L A T K K K K S Y H N P D S -
c N * Q Q E Y * Q R K R R R A I T I L I L -

TTAAAGATTTCAAAAATTCAGGTAAGAGAGATACATTCATTAAAATTCATATATTATAG
1081 -----+-----+-----+-----+-----+ 1140
AATTTCTAAAGTTTTTAAGGTCCATTCTCTATGTAAGTAATTTTAAGTATATAATATC

a L K I S K I P G K R D T F I K I H I L * -
b * R F Q K F Q V R E I H S L K F I Y Y S -
c K D F K N S R * E R Y I H * N S Y I I V -

TTTTTCATTTACAGCTGTTATTTTCTTTTATCTTAACAATATTTTTTGATTAGCTGGAA
1141 -----+-----+-----+-----+-----+ 1200
AAAAAGTAAAGTGTCGACAATAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT

a F F I S Q L L F S F I L T I F F D * L E -
b F S F H S C Y F L L S * Q Y F L I S W K -
c F H F T A V I F F Y L N N I F * L A G S -

GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGGTAACCTTAGCTTATTCACATTCAT
1201 -----+-----+-----+-----+-----+ 1260
CATTTTTTCATAGTTTATTTCTTCTCGGATCTGACTCCATTGAATCGAATAAGTGTAAGTA

a V K S I K * E K R * T E V T * L I H I H -
b * K V S N K R S A R L R * L S L F T F I -
c K K Y Q I R E A L D * G N L A Y S H S * -

AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTTAAAAA
1261 -----+-----+-----+-----+-----+ 1320
TCTAGCTGGAAGTATATAGGTTATGCTACTATTCTTTGTCGTCAGTAGGCAAAATTTTT

a R S T F I Y P I R * * G N S S H P F * K -
b D R P S Y I Q Y D D K E T A V I R F K N -
c I D L H I S N T M I R K Q Q S S V L K I -

TAGTGCTATGAGGACTAAATTTTTAGAGTCAAGAAATGGAGCCGAAATCTTAATCAAAAA
1321 -----+-----+-----+-----+-----+ 1380
ATCAGGATACTCCTGATTTAAAAATCTCAGTTCTTTACCTCGGCTTTAGAATTAGTTTTT

a * C Y E D * I F R V K K W S R N L N Q K -
b S A M R T K F L E S R N G A E I L I K K -
c V L * G L N F * S Q E M E P K S * S K R -

FIG. 12C



13/78

GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA
1381 -----+-----+-----+-----+-----+-----+ 1440
CTTAACGCAGCTATAACGTTTTCTTAGCTTGAGATTTAGAAAGCAATTATTCATAATGGT

a E L R R Y C K R I E L * I F R * * V L P -
b N C V D I A K E S N S K S F V N K Y Y Q -
c I A S I L Q K N R T L N L S L I S I T N -

ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA
1441 -----+-----+-----+-----+-----+-----+ 1500
TAGAACTAACTAATTCTCTAAGTCTCCGTTGACGTGTCTTCTAGTAATTTCTTTATTT

a I L I D C R D * R G N C T E D H * R N K -
b S C L I E E I D E A T A Q K I I K E I K -
c L D C L K R L T R Q L H R R S L K K * S -

GTAACCTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTTCAA
1501 -----+-----+-----+-----+-----+-----+ 1560
CATTGAAATAATTAATCTCTTATTTGATTAAATGATTATATCTCTAGTCGCTAGAAGTT

a V T F I N * R I N * I T N I E I S D L Q -
b * L L L I R E * T K L L I * R S A I F N -
c N F Y * L E N K L N Y * Y R D Q R S S I -

TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT
1561 -----+-----+-----+-----+-----+-----+ 1620
AACTGCTTTATTTTCGACTTGATTCAATCTGTTATTTTATGTTTGAACCAAGTTTAA

a L T K * K L N * S * T I K N T N L G Q N -
b C R N K S C T K V R Q * K I Q T L V K I -
c D E I K A E L K L D N K K Y K P W S K Y -

ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAATAAGGCAATAAATAAAATGA
1621 -----+-----+-----+-----+-----+-----+ 1680
TAACCTCCTTCCTTTTCTTCTGGTCAATCGTTTCTTTTATTTCCGTTATTTATTTTACT

a I E E G K E D Q L A K E K I R Q * I K C -
b L R K E K K T S * Q K K K * G N K * N E -
c C G R K R R P V S K R K N K A I N K M S -

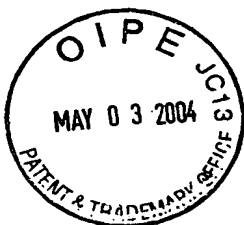
GTACAGAAGTGAAGAAATAAAAGATTTATTTTTCATAAATTTATTGAAAAGAGGGGTT
1681 -----+-----+-----+-----+-----+-----+ 1740
CATGTCTTCACTTCTTTATTTTCTAAATAAAAAAAGTTATTAAATAACTTTTCTCCCAA

a V Q K C R N K R F I F F N N L L K R G V -
b Y R S E E I K D L F F S I I Y C K E G F -
c T E V K K * K I Y F F Q * F I E K R G F -

TTGGGGTTTTGGGGTTTTGGGG
1741 -----+-----+-----+-----+-----+-----+ 1762
AACCCCAAACCCCAAACCC

a L G F W G F G -
b W G F G V L G -
c G V L G F W -

FIG. 12D



15/78

798 EKLINVSRENGFKFNMKK.LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846
: : : : : . | . | | | . : : : : . : : : : | . : : : :
577PGDEL RPSMQKLLQEKGKLG...TDFPYECIDEWTKNKTHVD 617
847 WIGISIDMKTLALMPNINLRIEGILCTLNLMQTKKASMWLKKKLSFLM 896
| . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . :
618 NIVILSDMMIAEGYSDINVRGSSIVNSI.....KKYKDEVN 653
897 NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMQCAKEYKD.HFKKNLAM 945
| | . : : . | : : : | : : . | : : . | : : . | : : . | : : .
654 PNIKIF...AVDLEGY...KCLNLGDEFNENNYIKIFGM 687
946 SSMIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFI 995
| . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . :
688 SDSI.....LKFISAKQGA.....NMVE 706
996 IFSTKKYIFNRVC 1008
: : . | . : : : :
707 VI..KNFALQKIG 717

FIG. 13B

132 LSTQKQYFFQDEWNQVRAMIGNEL.FRHLYTKYLIFQRTSE..GTLVQFC 178
: | . : | . : : : | | | : : : : : | . : : : :
1 MSRRNQ.....KKPQAPIGNETNLDFVLQNLLEVYKSQIEHYKTQQQQI 43
179 GNNVFDHLKVNDKFDKKQKGAADMNEPRCCSTCKYNVKNKEDHFLNNIN 228
: : . : | | . : : . : | . : | . : : : | . : : : :
44 KEEDLKLLKFKNQDQDGNsgnddddee.....NNSNKQOELLRRVN 84
229 VPNNWNMKSRTIRIFYCTHFNRRNQFFKKHEFVSNKNNISAMDRAQTFTN 278
: : : : | : : | | : | . : : : : : : : : :
85QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN 114
279 IFRFNIRKKLKDVKIEKIAYMLEKVKDFNFNYLTKSCPLPENWRERKQ 328
: . : : : . : : | . : : : | : : : : | . : . : : : | . : : : :
115 GLSEQQVKEEQLRTITTEEQVKYQNLVFNMDYQLDLNESGGHRRHRRETdy 164
329 KIENLINKTREESKYEEELFSYTTDNKCVTQFINE.FFYNILPKDFLTG 377
: . | . : : : : : : | . : : : : | . : : : : | . : : : :
165 DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK 200
378 RNRKNFQKKVKKYVELNKHელიHKNNLLEKINTREISWMQVETSAKHFFY 427
: | . : : : : : : | . : : : : : : | . : : : : : :
201 NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNSYQTVNID..... 242
428 FDHENIYVLWKLLRWI..FEDLVVSLIRCFYVTEQQKSYSKTYYYRKN 475
: : : | . : : : | | : : : | : : : | | : : : | . | . : : : :
243 VNFDNNLCILALLRFLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290
476 WDVIMKMSIADLKKETLAEVQEKEVEEWKSLGFAPGKLRLIPKKTTFRP 525
: . | . : . : : : : : | . : : : : : : : : : | . | . :
291 FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ 330
526 IMTFNKKIVNSDRKTTKLTNTNKLNSHMLKTLKNRMFKDPFGFAVFNY 575
: : | . : : : | . | . : : : | . : : : : | . : : : : | . | . :
331 VYSFSTDLKLVD..TNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENL 378

FIG. 14A



+

16/78

576 DDVMKKYEEFVCKWKQVGQPKLF.FATMDIEKCYDS..VNREK 615
: : : | | : : | | : : : | | : : : |
379 NVLLKKVKH ANLNLVSIPTQFNFDYFVNLQHLKLEFGLEPNILTKQK 426
516 LSTFL.KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKKEMK 657
| : : | | : : | | : : | | : : | | : : | | : : |
427 LENLLLSIKQSKNLKFLRLNFYTYVAQETSRKQILKQATTIKNLKNNKNQ 476
558 DYFRQKFQKIALEGGQYPTLFSVLEN..EQNDLNAKKTLLIVEAKQORNYFK 705
: : : | | : : | | : : | | : : | | : : | | : : |
477 EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY.. 520
706 KDNLLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSSFYATLEE 755
| | | | | : : | | : : | | : : | | : : | | : : |
521 .DSLHKLLIRSTNLKKFKLSYKYEMEKSKMDTFIDLKNI.....YETLNN 564
756 SSLGFLRDESMNPENPNVLLMRLTDDYLLITTQENNAVLFIKLINVSR 305
| : : | | : : | | : : | | : : | | : : | | : : |
565LKRCSVNISNPHGNISYELTN.....KDSTFYKFKLTNLNQE 500

806 ENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK 855
| : : | | : : | | : : | | : : | | : : | | : : |
601 LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDSLCKSIASCKNLQ 648
856 TLALMPNINLRIEGILCTLNLMQT..KKASMWLKK..KLKSFLMNNITH 901
: : : | : : | | : : | | : : | | : : | | : : | | : : |
649 NVNI.....IASLLYPNNIQKNPFNKPNLLFFKQFEQLKNLENVSINC 691
902 YFRKTI...TTEDFANKTLNKLFISSGGYKYMQCAKEYKDHFKKNLAMSSM 948
: : : | | : : : | | | | : : | | : : | | : : | | : : |
692 ILDQHILNSISEFLEKNKKIKAFILKRYLLQYYLDYTKLFKTLQQLPEL 741
949 IDLEVSKIISVT.....RAFFKYLVCNIKDT..IFGEEHY 982
: : : : | : : : | | : : | | : : | | : : | | : : |
742 NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESQTLQLIDF 791
983 PDFFLS TLKHFIEIFSTKKY IFNRVCMILKAKEAKLSDQCQSLIQ 1028
: : : | | : : | | : : | | : : | | : : | | : : | | : : |
792 DQNTVSDDSIKKILESISESKYHYLRLNPSQSSSLIKSENEEIQELLK 840

FIG. 14B

4 DIDLDDIENLLPNTFNKYSSSCSDKKGCKTLKSGSKSPSLTIPK..... 47
: : : | | : : : | | : : | | : : | | : : | | : : |
617 NVKSAKIESSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNP 666
48LQKQLEFYFSDANLYNDSFLRKLVLKSGEQORVE....IETLLM 86
: | | : : : | : : : | : : | | : : | | : : | | : : |
667 FNKPNLLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL 716

FIG. 15

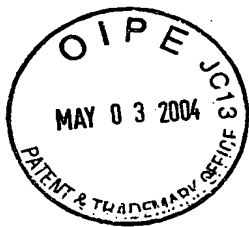
+



18/78

	Motif A	Motif B
Consensus	h--hDh---h--h	h---+QG---SP
telomerase p123	GPQLFFATMDIEKCYDSVNREKLSFLKTTKL-100-KFYKQTKGIPQGGLCVSSILSSFFYYATLEESSLGFL	
Dong (LINE)	KNRNLHCTYDDYKKAFDSIPHSLIQVLEIYKIN-28-RQIAIKKGIYQGDSLSPLWFCALNPLSHQLHNR	
al S.c. (groupII)	FGGSNWFREVDLKKCFDTISHDLIIKELKRYISD-26-HVPVGPRVCVQGAPTSPPALCNAVLLRLDRRLAGLA	
HIV-RT	LKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP-7-GIRYQYNVLPQGWKGSPPAIFQSSMTKIILEPFRKQN	
L8543.12	VLPELYFMKFDVKSCYDSIPRMECMRIKDALKN-68-KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK	
	Motif C	Motif D
Consensus	h--YhDDh	Gh-h---K
telomerase p123	-14-LMRLTDDYLLITTOENN-0-AVLFIKLINVSRENGFKFNMKKLT-23-QDYCDWIGISI	h-hLgh-h
Dong (LINE)	-16-HLIYMDDIKLYAKNDKE-0-MKKLIDTTTIFNSDISMQFGLDKCT-25-KCLYKYLGFQQ	
al S.c. (groupII)	-55-YVRYADDILIGVLGSKN-2-KIIRDLNFFLNS, LGLTINEEKTII-4-ETPARFLGYNI	
HIV-RT	-4-IYQYMDLLYVGSHEIG-1-HRTKIEELRQHLRWGLTTPDKKHQK-0-EPPFLWMGYEL	
L8543.12	-8-ILKLADDFLIISTDQQQ.....VINIKLAMGGFQKYNKANR-41-IRSKSKGIFR	

FIG. 17



19/78

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAACLLQEGSSYYQDKDERRYIITKALLEVAESDPEFICQLA
VYIRNELYIRTTTNYIVAFVHVHKNTPFIEKYFNKAVLLPNDL
LEVCEFAQVLYIFDATEFKNLYLDRILSQDIRKELTFRKCLQRC
VRSKFSEFNEYQLGKYCTESQRKKTMTFRYLSVTNKQKWDQTKKK
RKENLLTKLQAIKESEDKSKRETGDIMNVEDAIKALKPAVMKKI
AKRQONAMKKHMKAPKIPNSTLESKYLTFKDLIKFCHISEPKERV
YKILGKKYPKTEEEYKAAFQDSASAPFPNPELAGKRMKIEISKTW
ENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGVSDTT
HSIVINKICEPKAVENSKMFLQFFSAIEAVNEAVTKGFKAKKR
ENMNLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIG
KQYINSIELAIIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSSGA
KKYGSVRTCLECALVLGLMVKQRCESKFYIFSSPSSQCNKCYL
EVDLPGLDELPRSMQKLLQEKGLGGGTDFPYECIDEWTKNKTIV
DNIVILSDMMIAEGYSIDINVRGSSIVNSIKKYKDEVNPNIKIFA
VDLEGYKCLNLGDEFNENNYIKIFGMSDSILKFISAKQGGANM
VEVIKNFALQKIGQK

FIG. 20

MSRRNQKKPQAPIGNETNLDFVLQNLLEVYKSQIEHYKTQQQKIK
EEDLKLKFKNQDQDGNNGNDDDEENNSNKQOELLRRVNQIKQ
QVQLIKKVGSKVEKDLNLNEDENKKNGLSEQQVKEEQLRTITEE
QVKYQNLVFNMDYQLDLNESGGHRRHRRETDDYDEKWFESHDQ
KNYVSIYANQKTSYCWWLKDYFNKNYDHLNVSINRLETEAEFY
AFDDFSQTIKLTNNSYQTVNIDVNFDDNNLCILALLRFLLSLERF
NILNIRSSYTRNQYNFEKIGELLETFIAVVFSHRHLQGIHLQVP
CEAFQYLVNSSSQISVKDSQLQVYSFSTDLLKLVDTNKVQDYFKF
LQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNLVSI
TQFNFDYFVNLQHLKLEFGLPEPNILTKQKLENLLLSIKQSKNL
KFLRLNPHYTYVAQETSRKQILKQATTIKNLKNNKNQEETPETKD
ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDSLHKLII
RSTNLLKKFKLSYKYEMEKSKMDTFIDLKNIYETLNNLKRCSVNI
SNPHGNISYELTNKDSTFYKFKLTNLQELQHAKYTFKQNEFQFN
NVKSAKIBSSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPN
NIQKNPFPKNPNNLFFKQFEQLKNLENVSINCILDQHILNSISEF
LEKNKKIKAFILKRYLLQYYLDYTKLFKTLQQLPELNQVYINQ
QLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESQTLQLIDFD
QNTVSDDSIKKILESISESKYHHYLRNPSQSSSLIKSENEEIQ
ELLKACDEKGVLVKAYYKFPCLCLPTGTYDYNSDRW

FIG. 22

MKILFEFIQDKLDIDLQTNSTYKENLKC GFHNGLDEILTTCFAL
PNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKI
ARNEDVNNSLFCHSANVNVTLKGAAWKMFHSLVGTYAFVDLLI
NYTVIQFNGQFFTQIVGNRCNEPHLPPKWVQRSSSSSATAAQIK
QLTEPVTNKQFLHKLNNSSSFFPYSKILPSSSSSIKKLTDLREA
IFPTNLVKIPQRLKVRINLTQKLLKRHKRLNVYSILNSICPPL
EGTVLDLSHLRSRQSPKERVLFKIIIVILQKLLPQEMFGSKKNKGK
IIKLNLLLSLPLNGYLPFDSLLKKLRLKDFRWLFISDIWFTKH
NFENLNQLAICFISWLFRLPKIIQTFFCYCTEISSTVTIVYFR
HDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLNFNHNSKMR
IPKKSNNEFRIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEY
LRNKRPTSFTKIYSPTQIADRIKEFKQRLKKFNNVLPPELYFMK
FDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFNTNTGVL
KLFNVVNASRVKPYELYIDNVRTVHLSNQDVINVEMEIFKTA
LWVEDKCYIREDGLFQSSLSAPIVDLVYDDLLEFYSEFKASPS
QDTLILKLADDFLIISTDQQQVINIKKLAMGGFQKYNANRDK
ILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMNNFHRSKS
SKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISE
CYKSAFKDLSINVTQNMQFHSFLQRIIEMTVSGCPITKCDPLIE
YEVRFITLNGFLESLSNTSKFKDNIILLRKEIQHLQAYIYIYI
HIVN

FIG. 23



20/78

1 tcaatactat taattaataa ataaaaaaaaa gcaaaactaca aagaaaatgt caaggcgtaa
61 ctaaaaaaaa ccataggctc ctataggcaa tgaaacaaat cttgattttg tattacaaaa
121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaaga
181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
241 tgatgatgat gaagaaaaca actcaataaa ataataagaa ttattaagga gagtcaatta
301 gattaagtag caagtttaat tgataaaaaaa agttgggttct aaggtagaga aagatttgaa
361 tttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta
421 attaagaacg attactgaag aataggttaa gtattaaaat ttagtattta acatggacta
481 ccagtttagat ttaaatgaga gtgggtggcca tagaagacac agaagagaaa cagattatga
541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa ttacgccaa
601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaaca attatgatca
661 tcttaatgta agcattaaca gactagaaaa tgaagccgaa ttctatgctt ttctatgattt
721 ttcacaaaca atcaaaactta ctaataattc ttactagact gttaacatag acgttaattt
781 tgataataat ctctgtatac tcgcattgct tagattttta ttatcactag aaagattcaa
841 tattttgaaat ataagatctt cttatacaag aaattaatat aattttgaga aaattgggtga
901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cacttacaag gcattcattt
961 acaagttcct tgccaagcgt tctaattttt agttaactcc tcatcataaa tttagcgttaa
1021 agatagctaa ttataggtat actctttctc tacagactta aaattagttg acactaacaa
1081 agtccaagat tatttttaagt tcttataaga attccctcgt ttgactcatg taagctagta
1141 ggctatccca gttagtgcta ctaacgctgt agagaacctc aatgttttac ttaaaaaggt
1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctactt
1261 tgtaatttta taacattttga aattagagtt tggattagaa ccaaatattt tgacaaaaca
1321 aaagcttgaa aatctacttt tgagtataaa ataatacaaaa aatcttaaat ttttaagatt
1381 aaacttttac acctacgttg cttaaagaaa ctccagaaaa cagatattaa aacaagctac
1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaagatga
1501 aactccaagc gaaagcacia gtggtatgaa attttttgat catctttctg aattaaccga
1561 gcttgaagat ttcagcgtta acttgtaagc taccgaagaa atttatgata gcttgcacaa
1621 acttttgatt agatcaacaa attttaaagaa gttcaaatta agttacaaat atgaaatgga
1681 aaagagtaaa atggatacat tcatagactt taagaatatt tatgaaacct taaacaatct
1741 taaaagatgc tctgttaata tatcaaatcc tcatggaaac atttcttatg aactgacaaa
1801 taaagattct actttttata aattttaagct gaccttaaac taagaattat aacacgctaa
1861 gtatactttt aagtagaacg aattttaatt taataacggt aaaagtgcaa aaattgaatc
1921 ttcttcatta gaaagcttag aagatattga tagtctttgc aaatctattg cttcttgtaa
1981 aaattttacaa aatgttaata ttatcgccag ttgtctctat cccaacaata tttagaaaaa
2041 tcctttcaat aagcccaatc ttctattttt caagcaattt gaataattga aaaatttgga
2101 aaatgtatct atcaactgta ttcttgatca gcatataactt aattctattt cagaattctt
2161 agaaaagaat aaaaaataa aagcattcat tttgaaaaga tattatttat tacaatatta
2221 tcttgattat actaaattat ttaaaacact tcaatagtta cctgaattaa attaagttta
2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaaa
2341 ccacaagcaa aaagctttct atgaaccatt atgtgagttt atcaaagaat catcctaaac
2401 cctttagcta atagattttg accaaaacac tgtaagtgat gactctatta aaaagatttt
2461 agaactcata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttaatctag
2521 cagtttaatt aaatctgaaa acgaagaaat ttaagaactt ctcaaagctt gcgacgaaaa
2581 aggtgtttta gtaaaagcat actataaatt cctctatgt ttaccaactg gtacttatta
2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaaataaa tattaaatat
2701 tgaatatattc tttgcttatt atttgaataa tacatacaat agtcattttt agtgttttga
2761 atatatttta gttattttaat tcattatttt aagtaaataa ttattttttc atcatttttt
2821 aaaaaatcg

FIG. 21



21/78

Oxytricha
Euplotes

LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
LCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIG. 24

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTTCGCACTACCAAATTCAGAAAAATAGCATTACCATGCCTTCCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT
ATACAACAACGTACTAACATTTGGCTATAAAAATAGCTAGAAATGAAGATGTCAACAATAG
TCTTTTTTGGCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGAAAAAT
GTTCCACAGTTTGGTTCGGTACATACGCATTTCGTTGATTTATTGATCAATTATACAGTAAT
TCAATTTAATGGGCAGTTTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT
GCCGCCCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAAAC
ACTTACAGAACCAGTGACAAATAAACAAATTCCTTACACAAGCTCAATATAAATTCCTCTTC
TTTTTTTCTTATAGCAAGATCCTTCCTTCATCATCTATCAAAAAGCTAACTGACTT
GAGAGAAGCTATTTTTTCCCACAAATTTGGTTAAAATTCCTCAGAGACTAAAGGTACGAAT
TAATTTGACGCTGCAAAAGCTATTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT
GAATAGTATTTGCCCCACCATTTGGAAGGGACCGTATTGGACTTGTGCGATTTGAGTAGGCA
ATCACCAAAGGAACGAGTCTTGAAATTTATCATTTGTTATTTTACAGAAGTTATTACCCCA
AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATTTATCAAGAACTAAATCTTTTATT
AAGTTTACCCCTAAATGGCTATTTACCATTTGATAGTTTGTGAAAAAGTTAAGATTAA
GGATTTTTCGGTGGTTGTTTCATTTCTGATATTTGGTTACCAAGCACAAATTTTGAAAACTT
GAATCAATTGGCGATTTGTTTCATTTCTGGCTATTTAGACAACCTAATTCCTCAAAATTTAT
ACAGACTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA
TGATACTTGGAATAAACTTATCACCCTTTTATCGTAGAATATTTTAAAGACGTACTTAGT
CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA
AATGAGGATTATACCAAAAAAAGTAATAATGAGTTCAGGATTATTGCCATCCCATGCAG
AGGGGCAGACGAAGAAGAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC
CACTCAAAAAATTTTAGAATACCTAAGAAACAAAAGGCCGACTAGTTTTACTAAATATA
TTCTCCAACGCAATAGCTGACCGTATCAAAGAAATTTAAGCAGAGACTTTTAAAGAAATTT
TAATAATGTCTTACCAGAGCTTTATTTTCATGAAATTTGATGTCAAATCTTGCTATGATTC
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAATGGGTT
TTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTAAACAGCTTT
GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTTCAGGGCTCTAGTTTATC
TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTATAGCGAGTTTAAAGC
CAGTCCTAGCCAGGACACATTAATTTTAAACTGGCTGACGATTTCTTATAATATCAAC
AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAATATAA
TGCGAAAGCCAATAGAGACAAAATTTTAGCCGTAAGCTCCCAATCAGATGATGATACGGT
TATTCATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAACATTCAAG
CACAATGAATAATTTCCATATCCGTTTCAAAATCTAGTAAAGGGATATTTCAAGTTTAAAT
AGCGCTGTTTAACTAGAACTCTTTATAAAACAATTGACACAAATTTAAATTTCAACAAA
CACCGTTCTCATGCAAAATTGATCATGTTGTAAAGAACATTTTCGGAATGTTATAAATCTGC
TTTTAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTTCAATCGTTCTTACAACG
CATCATTGAAATGACAGTCAGCGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA
TGAGGTACGATTCACCATATTGAATGGATTTTTTGAAAGCCTATCTTCAAACACATCAAA
ATTTAAAGATAATATCATTCTTTTGAGAAAGGAAATTCAACTTGCAAGC

FIG. 26



23/78

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNM DYVVGARTFRREK
RAERLTSRVKALFSVLNYERA

FIG. 27

GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTC
TTTTCTTTTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAG
AGGGTGCACTGCGGGACGTGTCTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGC
CAGGCCCCGCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGC
TGCGGCCGATTGTGAACATGGACTACGTCTGTGGGAGCCAGAACGTTCCGCAGA
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTTCAGCGTGT
CAACTACGAGCGGGCGCG

FIG. 28

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
DSKPDEGVQFSSPKCSQSELIANVVKQMFDESFEERRNLLMKGFSMNHEDFRAMHVNGVQNDLVSTF
PNYLISILESKNWQLLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKKRKR
TIETSITQNK SARKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMLQWIFPRQFG
LINAQVKQLHKVIPLVSQSTVVPKRLLKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS
YSLKPNQVFAFLRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWL
VLGKRSNAKMCLSDFEKRKQIFAEFIYWLNSFIIPILQSFFYITESDLRNRVTYFRKDIWKLLCR
PFITSMKMEAFEKINENNVRMDTQKTTLP PAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTN
QTLRPVASILKHLIN EESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLM
FRIVKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVQVQLLSMKTSDTLFVDFVDY
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPOGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRVDDFLFITVNKKDAKKFLNLSLRGFEKHNFTSLEKTVINFENSNGIINNTFFNESKKRMPFFG
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN
IYRLGYSMCMRAQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLF
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFHRRRIAD

FIG. 29



24/78

ggtaaccgattacttcttcttataagtaattgcttctctcgaacgctcctaaatctctggaataattttacaaga
 actcaataaacaataccaagtcaaatcccaatgaaggtgttaagtagatcgatcgataataattttattttatcggttga
 ccaagtataaggacaaaaagaacaacttcttccccctaaagacttttactttacttaataatttacttttcaaatatttctg
 ggttcgcttacttttaactggttactgttttagtctactctctagccaacgggtgttctacccgctcatgtgat
 agctcttgagtagctcacagaaactcttacaatactctgatgagactatatttgattcattacagtcctgcatattc
 ttaacatggagccttacacttttagatgagtcacgctgcatgatggagtatttggtatcatccaagctttgaccttgaaaag
 gtbgataattattgcaaaatcatgcttagtgggttaactcgcgaagtttttggatgcttcacacgctctagcatg
 attgagatattcaaaaatttctatccactacaactcttaacgggtttttattttctatttctcattctcatgtgtt
 ccaatatgtatcatctgattaggcttttttcggttttactcctggaaatcgtaactttttcactattccccctaatga
 ataactaaaattggttcgcttataatgatagtagtagaagattgggtactctactcgtgaatgttattagtttaaa
 gatactttgcaaaacattttattagctatcattatataaaaaaaactcctataatataaaattataatcaatatttgcggtc
 actatttattaaaacgttatgacagtaggacactttgcatatataatagttatgcttaattggttacttgaacttgcat
 gaccgaacacattacccccaaagcagagattcttcgctttctagagaattcaatgtatattacattgtaccccttaaatgatt
 atgtacaacttggtttgaggggtcgccggcaagctcgtagcaataatatacggaacgcttgagaaagcgtatgtacaacg
 tcccttttctatttttcttcttctgactgtagtcgcttcgacagtaagccagatgaagggtgttcaattttcttctccaaa
 atgctcacagtcagaggtatataattttgtttgtttttttctctatcggatagctaataat999cagctaatagc
 gaattgtgtaaaaacagatgttcgatgaagttttgagcgtcgaaagaaatctactgataaagggttttccatggttaaggt
 attcctaattgtgaataattacctgcaattactgtttcaaaagattgtatttaaccgataaagaatcatgaagattttc
 gaccatgcatgtaaaacggagtagcaaaaatgattctgctttctacttttcttaattacattatatttactgtgagtcacaaa
 aattggcaacttttggtagaaatgtaaatccgggttaagatgtgctgcaacttgaacaagacagtaagtagtatcgg
 cagtatgccatgcatattctttatattccaaaggagattttttgaggtcttccaaatgacaaatfacccttcagattctg
 gcataccacttttttaaaaaataatgtgtttgagaaactgtgtcaaaaaaaagaaagcgaacccatgaaacatccattact
 caaaaataaaagcccgcaagagatttccctggaatagcatttcaatttagttagtttagcattttttacaggtcatccta
 taagaagtttaagcaaggtacctaactgttactctcactaaatttttagatctatatttttaactttacactctattt
 gtgattcggaacacagtagacatgtggcttcaatggatttttccaaaggcaatttggaacttataaagcattttcaagtgaag
 caattgcacaaagtgattccactggtagcacagtagacagttgtgccccaaacgctctcaaggtatacccttttaattga
 acaaacagcaaaaggactccatcgattttctctataaaagtttacaacattattgccccattatattgacacccacgattg
 atgaaaaaatcccttagttattccttaagccgaaccagggtgttgcgtttcttcgattccattctgttcgagtttctct
 aaattaatctgggttaaccgaagatatttgagataaattataaaaggtattgtataaaaattattaccactaacgatttt
 accagacctcgaaacttttcttgaaattatcgagatagcagttcttttagttttacattatttaagtagtaacataaaggtaa
 tatgcaaaattttttaccatttaatacaatcagattttcagaaattgaatggctagtccttggaanaaaggctcaaatgCG
 aaatgtgctttagtgatttttgaaaaacgcaagcaaaatattttgggaatttcactctactggctatfacaattcggttataat
 acctattttacaaacttttttttatatcatgaatcaagtagattacgaatcgaaactgtttatttttagaaaaagatttt
 ggaaactcttctgcccaccctttattacatcaatgaaatggaagcgttttgaaaaataaaacgaggtatttttaagatt
 ttttgcaaaaagctaaattttcagaacaatttaggattggaactcagaaaaacttttgcctccagcaggtttatctgctc
 tattacctaaagagaatacctttgctctatttagcgaatttagcaaaaaagatttcttaataaaggtatttaatttttgggtcat
 caatgtacttcttaactatttagcagattgggttcaaaaaaaaatgttagtcagtagcgaaccaaacttttacg
 acctgtggcatcgatactgaaacatttaattcaatgaagaaagtagtggtatttccattttaacttggagggttttacatgaagc

FIG. 30A

FIG. 30B

+



EST2 pep	FFYCTEISST VTIVYFRHDT WN-----KLIT P-----FIVE YPK-TYLVEN	40
Euplotes pep	FFYVTEQQKS YSKTYYYRKN IWDVI-MKMS IAD----LKK ETIA--EVQE	43
Trans of tetrahymen	-----KHKE GSQIFYRKP IWKLVSCLTI VKVRIQFSEK NKQMKNNFYQ	44
Consensus	FFY.TE..K. .S..YYRK. IW....KL.. ----F..K	50
EST2 pep	NVCRNHSY- ----- TLSNFNHSM RLTPKKSNE FRITAIPCRG	79
Euplotes pep	KEVEEWKSL ----- ---GFAPCKG RLTPKKT-- FRPIMTFNKK	78
Trans of tetrahymen	KIQLEENLE KVEEKLIPED SFQKYPQCKL RLTPKKS-- FRPIMTFLRK	92
Consensus	K...E..... -----F..GKL RLTPKK... FRPIMTF.RK	100
EST2 pep	ADEEEFTIYK ENHKNAIQPT QKILEYLRNK RPTSFTKIYS PTQIADRIKE	129
Euplotes pep	IVNSDRKTTK LTTNTKLLNS HMLKTLKN- -----RMFK -DPFGFAVEN	120
Trans of tetrahymen	DKQKNIK--- LNLNQILMDS QLVFRNLKD- -----ML-G -QKIGYSVFD	130
ConsensusK..K LN.N..L..S QL.L..LKN- -----...IG..VF.	150
EST2 pep	FKORLLKFN NVL----- PLYFMKFD VKSCYD	157
Euplotes pep	YD-DVMKKYE EFVCKWKQVH CPKLFATMD IEKCYD	155
Trans of tetrahymen	NK-QISEKFA QFIEKWKNG RPCLYYVTL- -----	158
Consensus	.K-....KCF. .F...KWK..G .P.LYF.T.D ...CYD	186

FIG. 31

+



27/78

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
 S-2: RQH LKR VQL RDV SEA EVR QHR EA
 S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
 A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
 A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

FIG. 32

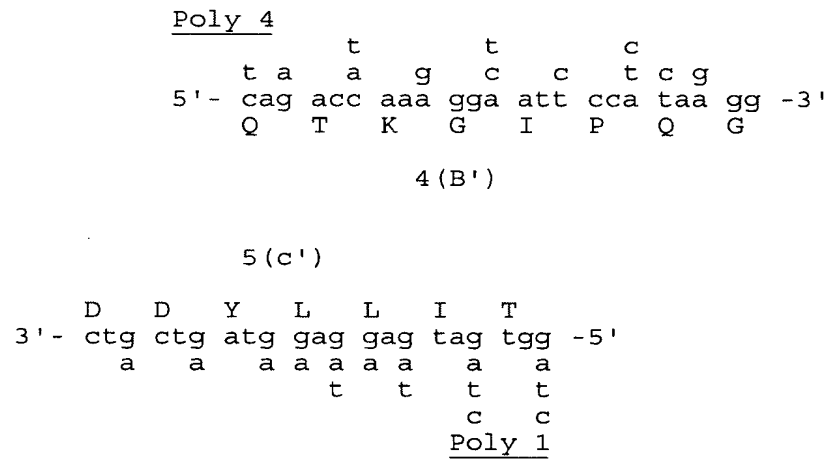


FIG. 34



28/78

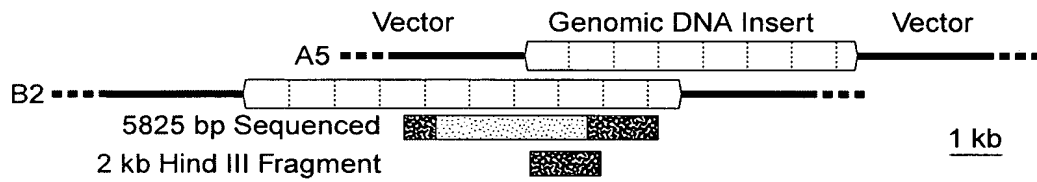


FIG. 33A

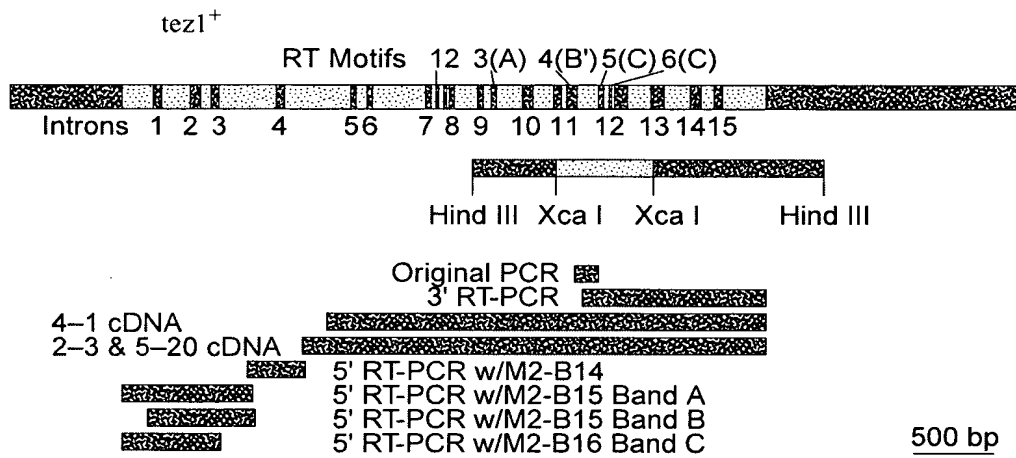


FIG. 33B

+



29/78

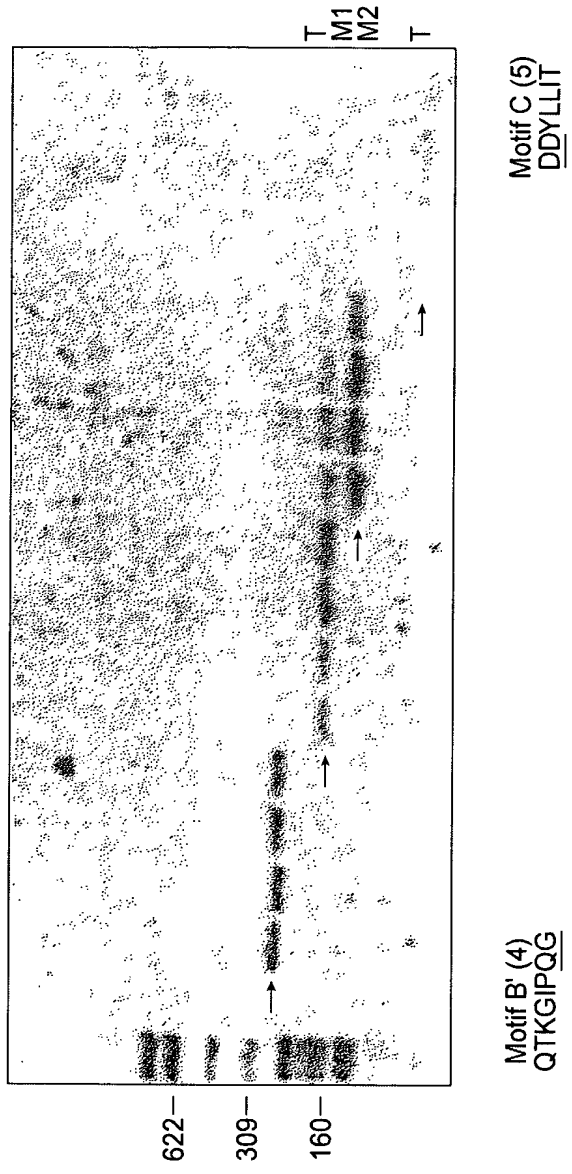


FIG. 35

+



30/78

Ot LCVSYILSSFYVANLEENALQFLRKESMDPEKPTNLLMRLLT
Ea_p123 KGIPOGLCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLMLRLTDDYLLIT
Sp_M2 SILSSFLCHFYMEDLLIDEYLSFTKKK-----GSVLLRVV
Sc_p103 DGLFQGSLSAPIVDLVYDDLLLEFYSEFKASPS-----QDTLILKLADDFLIIS
* . . . *

Q K V G I P Q G
caa aaa gtt ggt atc cct cag gg..... <----Actual Genomic Sequence.

Poly 4
t t c
t a a g c c t c g
cag acc aaa gga att cca taa gg ---->

ag acc aaa gga att cca tca ggc tca att ctg tca tct ttt ttg tgt cat ttc tat atg
tc tgg ttt cct taa ggt agt ccg agt taa gac agt aga aaa aac aca gta aag ata tac

K G I P S G S I L S S F L C H F Y M

FIG. 36A

+



GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT
E D L I D E Y L S F T K K K G S V L L R

GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

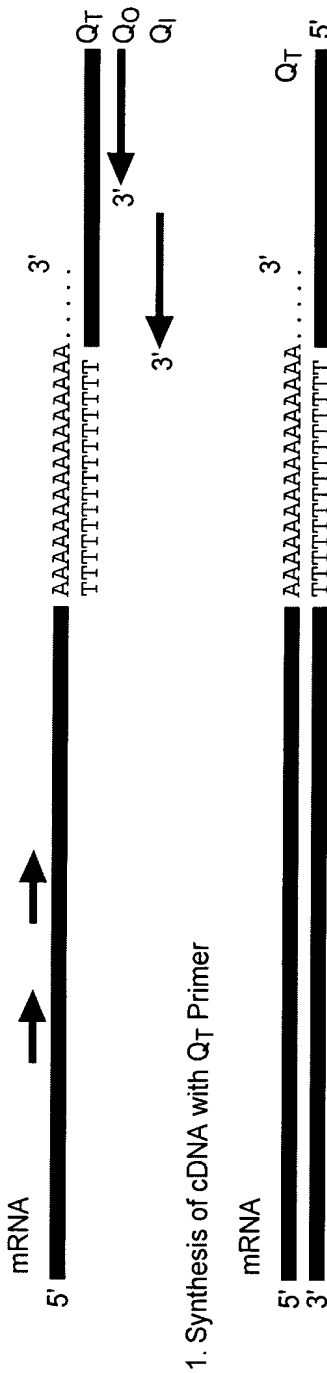
<----- ctg ctg atg gag gag tag tgg
a a a a a a a
t t t t t
c c
Poly 1

.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence
D D F L F I T

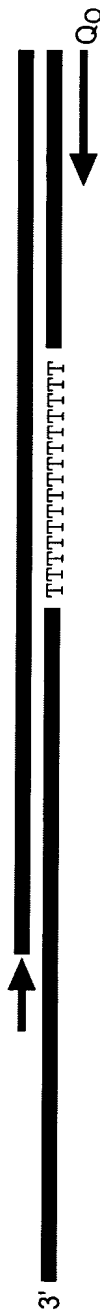
FIG. 36B

+

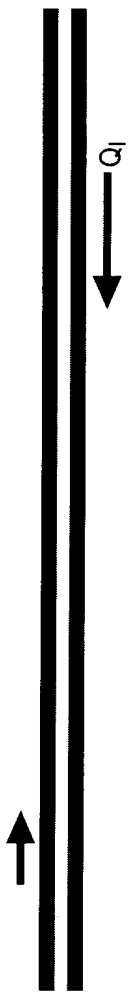
+



2. First Round PCR Using Outside Primer and Q_O Primer



3. Second Round PCR Using Inside Primer and Q_I Primer



4. Sequence Second Round PCR Products Using Inside Primer Q_I Primer

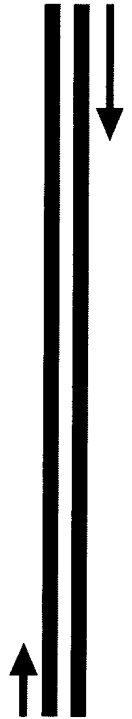


FIG. 37

+



+



33/78

- A. Genomic Libraries**
- Size Selected Libraries from P. Nurese
 - 3~4 kb
 - 5~7 kb
 - 7~8 kb
 - 11~12 kb
 - Libraries from J.A. Wise
 - Sau 3a Partial Digest
 - Hind III Partial Digest
- cDNA Libraries**
- GAD (Gal Activation Domain) Library
 - REP Library from R. Allshire
 - REP81ES Library (old)
 - REP81ES Library (new)
 - REP41ES Library

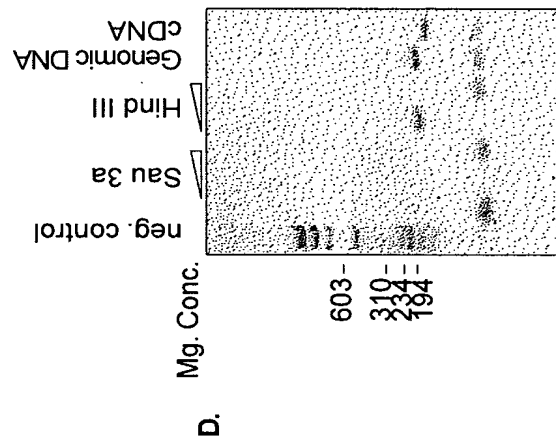
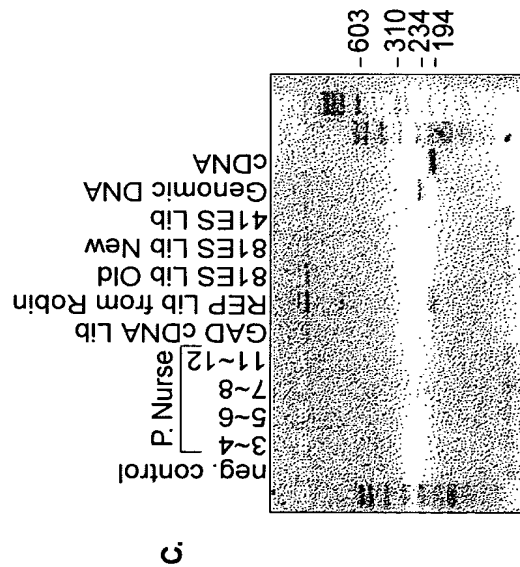
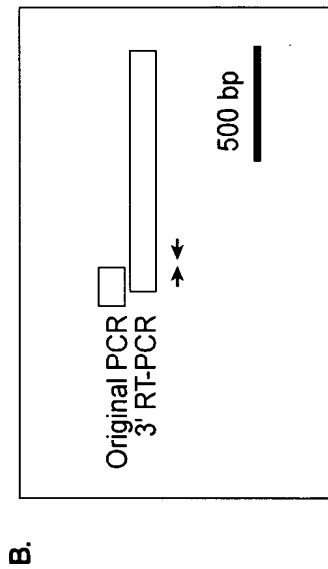


FIG. 38

+



34/78

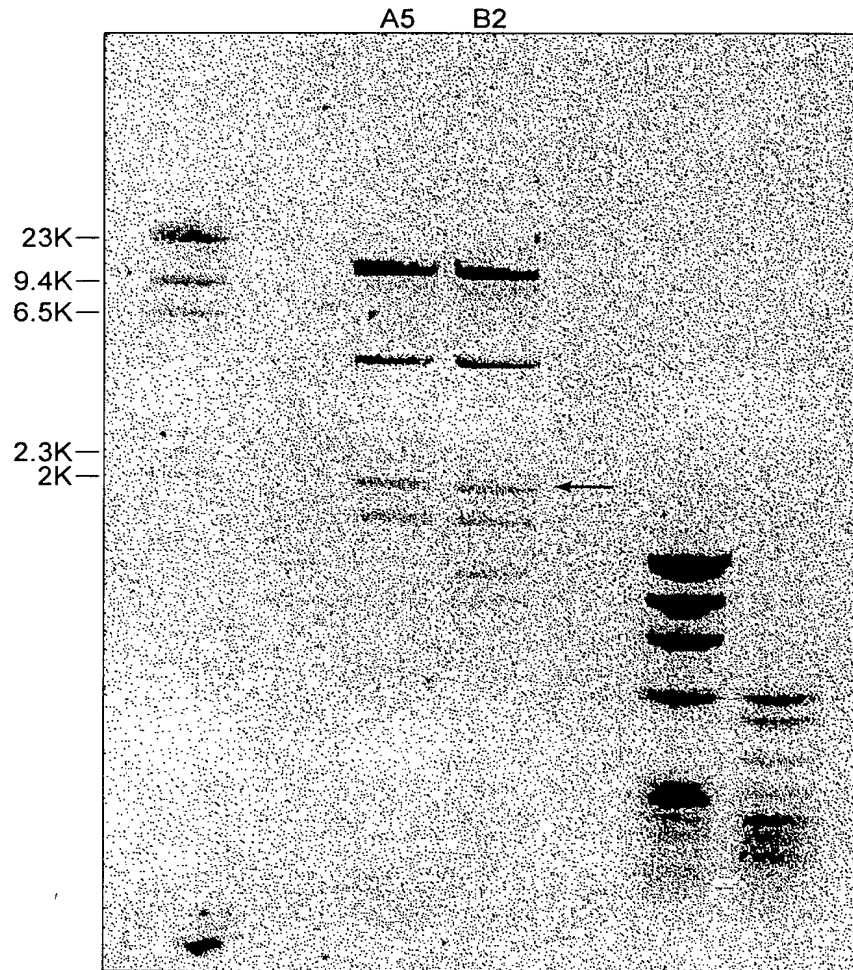


FIG. 39

+

+

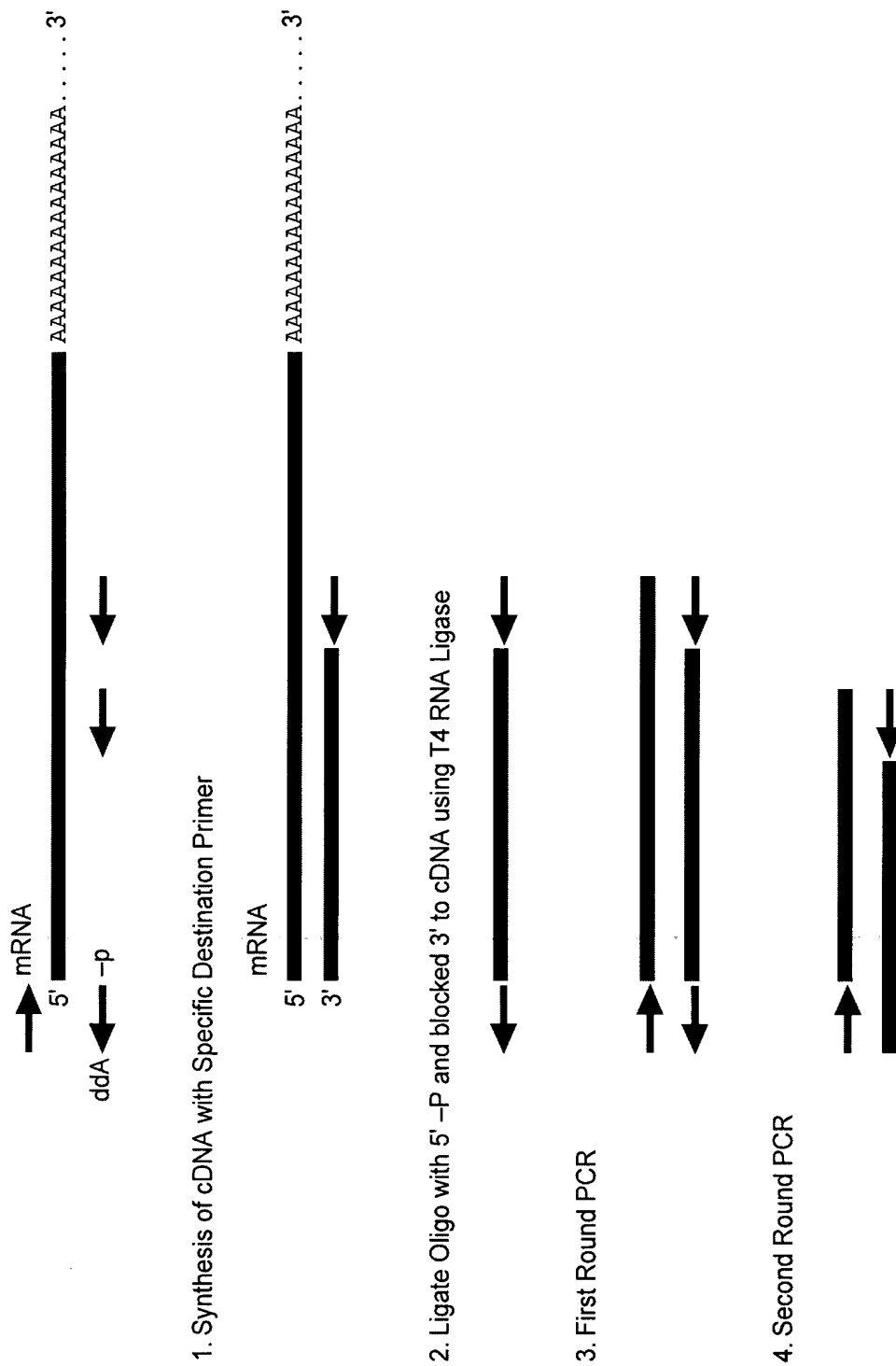


FIG. 40

+

+



Motif O		
S.p. Tezlp	(429) . WLYNSFIIPILQSPFFYITESSDLRNRVTYFRKDIW	... (35) ...
S.c. Est2p	(366) . WLFRLIPKIIQTFFYCTEISSTVT- IVYFRHDTW	... (35) ...
E.a. p123	(441) . WIFEDLVVSLIRCFYVTEQQKSYSTYYRKNIW	... (35) ...
* * * * *		
Motif 1 Motif 2 K		
p hh h K	hR h R	
S.p. Tezlp	AVIRLLPKK--NTFRLITN-LRKRF	... (61) ...
S.c. Est2p	SKMRIIPKKSNNFRIIAPCRGAD	... (62) ...
E.a. p123	GKRLIPKK--TFRPIMTFNKKIV	... (61) ...
* * * * *		
Motif 3 (A) AF		
h hDh GY h		
S.p. Tezlp	KKYFVRIDIKSCYDRIKQDLMFRIVK	... (89) ...
S.c. Est2p	ELYFMKFDVKSCYDSIPRMECMRILK	... (75) ...
E.a. p123	KLFFATMDIEKCYDSVNREKLSTFLK	... (107) ...
* * * * *		
Motif 4 (B')		
hPQG pP hh h		
S.p. Tezlp	YLQKVGIPQGSILSSFLCHFYMEDLIDEYLSF	... (6) ...
S.c. Est2p	YIREDGLFQGSLSAPIVDLVYDDLLIFYSEF	... (8) ...
E.a. p123	YKQTKGIPQGLCVSSILSSFYATLEESSLGF	... (14) ...
* * * * *		
Y Motif 5 (C)		
h F DDhhh		
S.p. Tezlp	VLLRVVDDFLFITVWNKKDAKKFLNLSLRGFEKHNFTSLEKTVINFENS	... (205)
S.c. Est2p	LILKLADDFLIISTDQQQVINIKKIAMGGFKYNAKANRDKILAVSSQS	... (173)
E.a. p123	LLMRLTDDYLLITTTQENNAVLFIKLIINVSRENGFKFNMKKLQTSFPLS	... (209)
* * * * *		
Motif 6 (D)		
Gh h cK h		

FIG. 41

+



37/78

A.

Sp_Tip1p	1	- - - - -	MTEHHTPKSRILRFL	ENQYVYLCT	24
Sc_Est2p	1	- - - - -	- - - - -	MKILFEF	7
Ea_p123	1	MEVDVDNQADNHG	HSALKTC	EIKEAKTL	33
Sp_Tip1p	25	LNDYVQLVLRGSP	ASSY	SNICER	57
Sc_Est2p	8	IQDKLDIDLQTN	- - - - -	ENLKCGH	35
Ea_p123	34	IQKVIRCRNQSQ	- - - - -	DLEDIKIFAQTN	61
Sp_Tip1p	58	IFLHSTVVGFD	SKPDEGVQFSSPKC	SQS	90
Sc_Est2p	36	EILTTCFALPNS	R-KIALPCLPGDL	SHKAV	67
Ea_p123	62	I VATPRDYNEED	FKVIA	RKEVFSTGLMIEL	94
Sp_Tip1p	91	VVKQMFDES	FERRR	- NLLMKGF	122
Sc_Est2p	68	CIYLLTGELYN	- - - - -	NVLTFGYK	93
Ea_p123	95	CLVELLS	SSSDVSDRQKL	QCFGFQLKGNQ	122
Sp_Tip1p	123	VNGVQNDLVST	FPN	YLSILESKN	155
Sc_Est2p	94	- - - VNN	SLFCHSAN	VNVTTLLKGA	123
Ea_p123	123	- - - LAKTHLLT	ALSTQKQYFFQDEWN	QVRAMIG	152
Sp_Tip1p	156	SDAMHYLLSKGS	IF	EALPNDNYLQIS	188
Sc_Est2p	124	TYAFVDLLIN	YTVIQFN	- GQFF	155
Ea_p123	153	NELFRHLYTKY	LIFQRT	SEGTLVQFCGNN	185
Sp_Tip1p	189	NVFEETVSKKR	KRTI	ETSI	218
Sc_Est2p	156	HLPPKWVQ	- - - R	SSSSATAAQ	183
Ea_p123	186	LK VNDKFDK	- KQKGGAA	ADMNEPRCCSTCKYN	217

FIG. 42A

+



A.

Sp_Tip1p	219	WNSISISRFSIF	YR	SS	Y	K	K	F	K	Q	D	L	Y	F	N	L	H	S	I	C	D	251		
Sc_Est2p	184	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	200		
Ea_p123	218	N	E	K	-	-	D	H	F	L	N	I	N	V	P	N	W	N	M	K	S	R	248	
Sp_Tip1p	252	R	N	T	V	H	M	W	L	Q	W	I	F	P	R	Q	F	L	I	N	A	F	284	
Sc_Est2p	201	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	223	
Ea_p123	249	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	275	
Sp_Tip1p	285	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	313	
Sc_Est2p	224	T	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	252	
Ea_p123	276	F	T	N	I	F	R	F	N	R	I	R	K	K	L	K	D	K	V	I	E	K	I	308
Sp_Tip1p	314	L	S	K	V	N	H	Y	C	P	Y	I	D	-	T	H	D	D	E	K	I	L	S	342
Sc_Est2p	253	Y	V	S	I	L	N	S	I	C	P	L	E	G	T	V	L	D	L	S	H	L	S	282
Ea_p123	309	F	N	Y	Y	L	T	K	S	C	P	L	P	E	N	W	R	E	R	K	Q	I	E	341
Sp_Tip1p	343	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	359	
Sc_Est2p	283	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	299	
Ea_p123	342	S	K	Y	Y	E	E	L	F	S	Y	T	T	D	N	K	C	V	T	Q	F	I	N	374
Sp_Tip1p	360	W	G	N	Q	R	I	F	E	I	I	L	K	D	L	E	T	F	L	K	L	S	R	392
Sc_Est2p	300	F	G	S	K	K	N	K	G	K	I	I	K	N	L	L	L	S	L	P	N	G	Y	332
Ea_p123	375	L	T	G	-	R	N	R	K	N	F	Q	K	K	V	K	Y	V	E	L	N	K	H	406
Sp_Tip1p	393	N	I	K	I	S	E	I	E	W	L	V	L	G	K	R	S	N	A	K	M	C	L	425
Sc_Est2p	333	K	L	R	L	K	D	F	R	W	L	F	I	S	-	-	-	-	-	-	-	-	-	362
Ea_p123	407	K	I	N	T	R	E	I	S	W	M	Q	V	E	T	S	-	A	K	H	F	Y	F	437

FIG. 42B

+

39/78

A.	Sp_Tip1p	426	E	F	I	Y	W	L	Y	N	S	F	I	I	P	I	L	Q	S	F	F	Y	I	T	E	S	S	D	L	R	N	R	T	V	Y	458
	Sc_Est2p	363	C	F	I	S	W	L	F	R	Q	L	I	P	K	I	I	Q	T	F	F	Y	C	T	E	I	S	S	T	V	T	-	I	V	Y	394
	Ea_p123	438	K	L	L	R	W	I	F	E	D	L	V	V	S	L	I	R	C	F	F	Y	V	T	E	Q	Q	K	S	Y	S	K	T	Y	Y	470
	Sp_Tip1p	459	F	R	K	D	I	W	K	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	N	N	V	R	M	D	491
	Sc_Est2p	395	F	R	H	D	T	W	N	K	L	I	T	P	F	I	V	E	Y	E	K	T	Y	L	V	E	N	N	V	C	R	N	H	N	S	427
	Ea_p123	471	Y	R	K	N	I	W	D	V	I	M	K	M	S	I	A	D	L	K	K	E	T	L	A	E	V	Q	E	K	E	V	E	E	W	503
	Sp_Tip1p	492	T	Q	K	T	T	L	P	A	V	I	R	L	L	P	K	K	-	-	N	T	F	R	L	I	T	N	L	R	K	R	F	L	522	
	Sc_Est2p	428	Y	T	L	S	N	F	N	H	S	K	M	R	I	I	P	K	K	S	N	N	E	F	R	I	I	A	I	P	C	R	G	A	D	460
	Ea_p123	504	K	K	S	L	G	F	A	P	G	K	L	R	L	I	P	K	K	-	-	T	T	F	R	P	I	M	T	F	N	K	K	I	V	534
	Sp_Tip1p	523	I	K	M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V	A	S	I	L	K	H	L	I	N	E	-	-	-	552
	Sc_Est2p	461	E	E	E	-	-	F	T	I	Y	K	E	N	H	K	N	A	I	Q	P	T	Q	K	I	L	E	Y	L	R	N	K	R	P	T	491
	Ea_p123	535	N	S	D	-	-	R	K	T	T	K	L	T	T	N	T	K	L	L	N	S	H	L	M	L	K	T	L	K	N	R	-	M	F	564
	Sp_Tip1p	553	E	S	S	G	I	P	F	N	L	E	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G	R	-	K	K	584
	Sc_Est2p	492	S	F	T	K	I	Y	S	P	T	Q	I	A	D	R	I	K	E	F	K	Q	R	L	L	K	K	F	N	N	V	L	P	E	L	524
	Ea_p123	565	K	D	P	F	G	F	A	V	F	N	Y	D	D	V	M	K	K	Y	E	E	F	V	C	K	W	K	Q	V	G	P	K	L	597	
	Sp_Tip1p	585	Y	F	V	R	I	D	I	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	K	L	K	D	P	E	-	F	616
	Sc_Est2p	525	Y	F	M	K	F	D	V	K	S	C	Y	D	S	I	P	R	M	E	C	M	R	I	L	K	D	A	L	K	N	E	N	G	F	557
	Ea_p123	598	F	F	A	T	M	D	I	E	K	C	Y	D	S	V	N	R	E	K	L	S	T	F	L	K	T	T	K	L	L	S	S	D	F	630
	Sp_Tip1p	617	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	T	K	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	634
	Sc_Est2p	558	F	V	R	S	Q	Y	F	F	N	T	N	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	570	
	Ea_p123	631	W	I	M	T	A	Q	I	L	K	R	K	N	I	V	I	D	S	K	N	F	R	K	K	E	M	K	D	Y	F	R	Q	K	663	

FIG. 42C



40/78

A.	Sp_Tip1p	635	F	V	S	E	A	F	S	Y	F	D	M	V	P	F	E	K	V	V	Q	L	L	S	-	-	M	K	T	S	D	T	L	F	V	665	
	Sc_Est2p	571	-	-	-	-	-	-	-	-	-	-	-	V	L	K	L	F	N	V	V	N	A	S	R	-	-	V	P	K	P	Y	E	L	Y	I	591
	Ea_p123	664	F	Q	K	I	A	L	E	G	G	Q	Y	P	T	L	F	S	V	L	E	N	E	Q	N	D	L	N	A	K	K	T	L	I	V	696	
	Sp_Tip1p	666	D	F	V	D	Y	W	T	K	S	S	S	E	I	F	K	M	L	K	E	H	L	S	G	H	I	V	K	I	G	N	S	Q	Y	698	
	Sc_Est2p	592	D	N	V	R	T	V	H	L	S	N	Q	D	V	I	N	V	V	E	M	E	I	F	K	T	A	L	W	V	E	D	K	C	Y	624	
	Ea_p123	697	E	A	K	Q	R	N	Y	F	K	K	D	N	L	L	Q	P	V	I	N	I	C	Q	Y	N	Y	I	N	F	N	G	K	F	Y	729	
	Sp_Tip1p	699	L	Q	K	V	G	I	P	Q	G	S	I	L	S	S	F	L	C	H	F	Y	M	E	D	L	I	D	E	Y	L	S	F	T	K	731	
	Sc_Est2p	625	I	R	E	D	G	L	F	Q	G	S	S	L	S	A	P	I	V	D	L	V	Y	D	D	L	L	E	F	Y	S	E	F	K	A	657	
	Ea_p123	730	K	Q	T	K	G	I	P	Q	G	L	C	V	S	S	I	L	S	S	F	Y	A	T	L	E	E	S	S	L	G	F	L	R	762		
	Sp_Tip1p	732	K	K	G	-	-	-	-	-	-	-	-	-	-	S	V	L	L	R	V	V	D	D	F	L	F	I	T	V	N	K	K	D	A	K	756
	Sc_Est2p	658	S	P	S	Q	D	-	-	-	-	-	-	-	-	T	L	I	L	K	L	A	D	D	F	L	I	S	T	D	Q	Q	Q	V	I	N	684
	Ea_p123	763	D	E	S	M	N	P	E	N	P	N	V	N	L	L	M	R	L	T	D	D	Y	L	L	I	T	T	Q	E	N	N	A	V	L	795	
	Sp_Tip1p	757	F	L	N	L	S	L	R	G	F	E	K	H	N	F	S	T	S	L	E	K	T	V	I	N	F	E	N	S	N	G	-	-	-	786	
	Sc_Est2p	685	I	K	K	L	A	M	G	G	F	Q	K	Y	N	A	K	A	N	R	D	K	I	L	A	V	S	S	Q	S	D	-	-	-	-	713	
	Ea_p123	796	F	I	E	K	L	I	N	V	S	R	E	N	G	F	K	F	N	M	K	L	Q	T	S	F	P	L	S	P	S	K	F	A	828		
	Sp_Tip1p	787	-	-	-	I	I	N	N	T	F	F	N	E	S	K	K	R	M	P	F	F	G	F	S	V	N	M	R	S	L	D	T	L	L	816	
	Sc_Est2p	714	-	-	-	D	D	T	V	I	Q	F	C	A	-	-	M	H	I	F	V	K	E	L	E	V	W	K	H	S	S	T	M		739		
	Ea_p123	829	K	Y	G	M	D	S	V	E	E	Q	N	I	V	Q	D	Y	C	D	W	I	G	I	S	I	D	M	K	T	L	A	L	M	P	861	
	Sp_Tip1p	817	A	C	P	K	I	D	E	A	L	F	N	S	T	S	V	E	L	T	K	H	M	G	K	S	F	F	Y	K	I	L	R	S	S	849	
	Sc_Est2p	740	N	N	F	H	I	R	S	K	S	S	K	G	I	F	R	S	L	I	A	L	F	N	T	R	I	S	Y	K	T	I	D	T	N	772	
	Ea_p123	862	N	I	N	L	R	I	E	G	I	L	C	T	L	N	L	N	M	Q	T	K	K	A	S	M	W	L	K	K	L	K	S	F	894		

FIG. 42D

+



A.	Sp_Tip1p	850	L	A	S	F	A	Q	V	F	I	D	I	T	H	N	S	K	F	N	S	C	C	N	I	Y	R	L	G	Y	S	M	C	M	R	882	
	Sc_Est2p	773	L	N	S	T	N	T	V	L	M	Q	I	D	H	V	V	K	N	I	S	E	C	-	-	-	-	-	-	-	-	-	-	-	-	-	793
	Ea_p123	895	L	M	N	N	I	T	H	Y	F	R	K	T	I	T	E	D	F	A	N	K	T	L	N	K	L	F	I	S	G	G	Y	K	927		
	Sp_Tip1p	883	A	Q	A	Y	L	K	R	M	K	D	I	F	I	P	Q	R	M	F	I	T	D	L	L	N	V	I	G	R	K	I	W	K	K	915	
	Sc_Est2p	794	-	-	-	Y	K	S	A	F	K	D	L	S	I	N	-	-	V	T	Q	N	M	Q	F	H	S	F	L	Q	R	I	I	E	M	821	
	Ea_p123	928	Y	M	Q	C	A	K	E	Y	K	D	H	F	K	K	N	L	A	M	S	S	M	I	D	L	E	V	S	K	I	I	Y	S	V	960	
	Sp_Tip1p	916	L	A	E	I	L	G	Y	T	S	R	R	F	L	S	S	A	E	V	K	W	L	F	C	L	G	M	R	D	G	L	K	P	S	948	
	Sc_Est2p	822	T	V	S	G	C	P	I	T	K	C	D	P	L	I	E	Y	E	V	R	F	T	I	L	N	G	F	L	E	S	L	S	S	N	854	
	Ea_p123	961	T	R	A	F	F	K	Y	L	V	C	N	I	K	D	T	I	F	G	E	E	H	Y	P	D	F	F	L	S	T	L	K	H	F	993	
	Sp_Tip1p	949	F	K	Y	H	P	C	F	E	Q	L	I	Y	Q	F	Q	S	L	T	D	L	I	K	P	L	R	P	V	L	R	Q	V	L	F	981	
	Sc_Est2p	855	T	S	-	-	-	-	-	-	-	-	-	-	-	K	F	K	D	N	I	L	L	R	K	E	I	Q	H	L	Q	A	Y	I	Y	877	
	Ea_p123	994	I	E	I	F	S	-	-	-	-	-	-	-	-	T	K	K	Y	I	F	N	R	V	C	M	I	L	K	A	K	E	A	K	L	1023	
	Sp_Tip1p	982	L	H	R	R	I	A	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	988		
	Sc_Est2p	878	I	Y	I	H	I	V	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	884			
	Ea_p123	1024	Q	S	L	I	Q	Y	D	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1031		

FIG. 42E

+



42/78

B.

Sp_Tip1p	1	- - - - -	MTEHHTPKSRILRFL	ENQYVYLCT	24
Sc_Est2p	1	- - - - -	- - - - -	MKILFEF	7
Ea_p123	1	MEVDVDNQADNHG	HSALKTCEEIKEAK	TLYSW	33
Sp_Tip1p	25	LNDYVQLVLRGSPA	SSY	SNICERLRSDVQTSFS	57
Sc_Est2p	8	IQDKLDIDLQTN	- - - - -	ENLKCGHFNGLD	35
Ea_p123	34	IQKVIRCRNQSQ	- - - - -	DLEDIKIFAQTN	61
Sp_Tip1p	58	IFLHSTVVGFDSK	PDEGVQFSSPKCSQSEL	IAN	90
Sc_Est2p	36	EILTTTCFALPNSR	- KIALPCLPGDLSHKA	VIDH	67
Ea_p123	62	I VATPRDYNEED	FKVIARKEVFSTGLMIEL	LDK	94
Sp_Tip1p	91	VVKQMFDESFERR	- NLLMK	GFSMNHEDFRAMH	122
Sc_Est2p	68	CIIYLLTGELYN	- - - - -	GVYKIAARNED	93
Ea_p123	95	CLVELLSSSDVSDR	QKLQCF	GFGQLKGNQ	122
Sp_Tip1p	123	VNGVQNDLVSTFP	NYLISILESKN	WQLLEIIG	155
Sc_Est2p	94	- - - VNNSLFCH	SANVNVTLLKGA	AWKMFHSLVG	123
Ea_p123	123	- - - LAKTHLLT	ALSTQKQYFFQDE	WNQVRAMIG	152
Sp_Tip1p	156	SDAMHY	L LSKGSI	FEALPNDNYLQ	ISG
Sc_Est2p	124	TYAFVD	LLIN	YTVIQFN	- GQFF
Ea_p123	153	NELFRH	LYTKYL	IFQRTSEGT	LVQFCG
Sp_Tip1p	189	NVFEETVSKKRKR	TIETSI	TQN	- - - KSARKE
Sc_Est2p	156	HLPPKWVQ	- - - RSSSS	SATAAQI	- - - KQLTEP
Ea_p123	186	LKVNDFDK	- KQKGG	AADMNEPRCC	STCKYNVK

FIG. 42F



43/78

B.

Sp_Tip1p	219	WNSISISRF	SI	FYR	SS	YK	KFK	QDL	YF	NL	HS	IC	D	251
Sc_Est2p	184	N	-	-	-	-	-	-	-	-	-	-	-	200
Ea_p123	218	NEK	-	-	DH	FL	NN	IN	VP	NN	MM	KS	RT	248
Sp_Tip1p	252	RNTVHMWL	QW	IF	PR	QF	GL	IN	AF	VQ	VK	QH	KV	284
Sc_Est2p	201	-	-	-	-	-	-	-	-	-	-	-	-	223
Ea_p123	249	R	-	-	-	-	-	-	-	-	-	-	-	275
Sp_Tip1p	285	VS	-	-	-	-	-	-	-	-	-	-	-	313
Sc_Est2p	224	TN	-	-	-	-	-	-	-	-	-	-	-	252
Ea_p123	276	FTN	IF	R	F	N	R	I	R	K	K	L	K	308
Sp_Tip1p	314	LSK	V	N	H	Y	C	P	Y	I	D	-	-	342
Sc_Est2p	253	YV	S	I	L	N	S	I	C	P	L	E	G	282
Ea_p123	309	F	N	Y	L	T	K	S	C	P	L	P	E	341
Sp_Tip1p	343	-	-	-	-	-	-	-	-	-	-	-	-	359
Sc_Est2p	283	-	-	-	-	-	-	-	-	-	-	-	-	299
Ea_p123	342	SK	Y	E	E	L	F	S	Y	T	T	D	N	374
Sp_Tip1p	360	WGN	Q	R	I	F	E	I	I	L	K	D	L	392
Sc_Est2p	300	FG	S	K	K	N	K	G	K	I	I	K	N	332
Ea_p123	375	L	T	G	-	R	N	R	K	N	F	Q	K	406
Sp_Tip1p	393	N	I	K	I	S	E	I	E	W	L	V	L	425
Sc_Est2p	333	K	L	R	L	K	D	F	R	W	L	F	I	362
Ea_p123	407	K	I	N	T	R	E	I	S	W	M	Q	V	437

FIG. 42G



44/78

B.

Sp_Tip1p	426	EF	I	Y	W	L	Y	N	S	F	I	P	I	L	Q	S	F	F	Y	I	T	E	S	S	D	L	R	N	R	T	V	Y	458		
Sc_Est2p	363	C	F	I	S	W	L	F	R	Q	L	I	P	K	I	Q	T	F	F	Y	C	T	E	I	S	S	T	V	T	-	I	V	Y	394	
Ea_p123	438	K	L	L	R	W	I	F	E	D	L	V	S	L	I	R	C	F	F	Y	V	T	E	Q	Q	K	S	Y	S	K	T	Y	Y	470	
Sp_Tip1p	459	F	R	K	D	I	W	K	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	N	N	V	R	M	D	491
Sc_Est2p	395	F	R	H	D	T	W	N	K	L	I	T	P	F	I	V	E	Y	F	K	T	Y	L	V	E	N	N	V	C	R	N	H	N	S	427
Ea_p123	471	Y	R	K	N	I	W	D	V	I	M	K	M	S	I	A	D	L	K	K	E	T	L	A	E	V	Q	E	K	E	V	E	E	W	503
Sp_Tip1p	492	T	Q	K	T	T	L	P	P	A	V	I	R	L	L	P	K	K	-	-	N	T	F	R	L	I	T	N	L	R	K	R	F	L	522
Sc_Est2p	428	Y	T	L	S	N	F	N	H	S	K	M	R	I	I	P	K	K	S	N	N	E	F	R	I	I	A	I	P	C	R	G	A	D	460
Ea_p123	504	K	K	S	L	G	F	A	P	G	K	L	R	L	I	P	K	K	-	-	T	T	F	R	P	I	M	T	F	N	K	K	I	V	534
Sp_Tip1p	523	I	K	M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V	A	S	I	L	K	H	L	I	N	E	-	-	-	552
Sc_Est2p	461	E	E	E	-	-	F	T	I	Y	K	E	N	H	K	N	A	I	Q	P	T	Q	K	I	L	E	Y	L	R	N	K	R	P	T	491
Ea_p123	535	N	S	D	-	-	R	K	T	T	K	L	T	T	N	T	K	L	L	N	S	H	L	M	L	K	T	L	K	N	R	-	M	F	564
Sp_Tip1p	553	E	S	S	G	I	P	F	N	L	E	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G	R	-	K	K	584
Sc_Est2p	492	S	F	T	K	I	Y	S	P	T	Q	I	A	D	R	I	K	E	F	K	Q	R	L	L	K	K	F	N	N	V	L	P	E	L	524
Ea_p123	565	K	D	P	F	G	F	A	V	F	N	Y	D	D	V	M	K	K	Y	E	E	F	V	C	K	W	K	Q	V	G	Q	P	K	L	597
Sp_Tip1p	585	Y	F	V	R	I	D	I	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	K	L	K	D	P	E	-	F	616
Sc_Est2p	525	Y	F	M	K	F	D	V	K	S	C	Y	D	S	I	P	R	M	E	C	M	R	I	L	K	D	A	L	K	N	E	N	G	F	557
Ea_p123	598	F	F	A	T	M	D	I	E	K	C	Y	D	S	V	N	R	E	K	L	S	T	F	L	K	T	T	K	L	L	S	S	D	F	630
Sp_Tip1p	617	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	T	K	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	634
Sc_Est2p	558	F	V	R	S	Q	Y	F	F	N	T	N	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	570	
Ea_p123	631	W	I	M	T	A	Q	I	L	K	R	K	N	I	V	I	D	S	K	N	F	R	K	K	E	M	K	D	Y	F	R	Q	K	663	

FIG. 42H



45/78

B.

Sp_Tip1p	635	FVSEAFSYFDMVPFEK	V	VQLLS	-	-	MKTSDTL	FV	665
Sc_Est2p	571	-	-	-	-	-	VLKLFNVNASR	-	591
Ea_p123	664	FQKIALEGGQYPTLFS	V	LENEQN	D	L	NAKKT	L	696
Sp_Tip1p	666	DFVDYWTKSSSE	I	FKMLKEHL	SGH	I	VKIGNSQ	Y	698
Sc_Est2p	592	DNVRTVHLSNQDV	I	VVEME	I	FKTALW	VEDKCY		624
Ea_p123	697	EAKQRNYFKKDN	L	LQPV	I	ICQY	YIN	FN	729
Sp_Tip1p	699	LQKVG	I	PQGS	I	LSFLCHFY	MED	L	731
Sc_Est2p	625	IREDGLFQGS	SL	SAPI	V	DLVYDD	L	LEFYSE	762
Ea_p123	730	KQTKG	I	PQGLCV	S	I	LSFYAT	L	762
Sp_Tip1p	732	KKG	-	-	-	-	SVLLRV	VDD	756
Sc_Est2p	658	SPSQD	-	-	-	-	TLILKL	ADD	684
Ea_p123	763	DESMNPEN	P	NVNL	ML	RL	TD	Y	795
Sp_Tip1p	757	FLNLSLRGFEKH	N	FST	S	LEK	TV	I	786
Sc_Est2p	685	IKKLAMGGFQ	KYN	AKAN	RD	KI	LA	V	713
Ea_p123	796	FIEKLINVSRE	NGFK	FN	MM	KK	LQ	T	828
Sp_Tip1p	787	-	-	-	I	INNTFF	NESK	KRM	816
Sc_Est2p	714	-	-	-	DD	TVIQ	FCA	-	739
Ea_p123	829	KYGMDSV	EEQN	I	VQD	YCD	WIG	I	861
Sp_Tip1p	817	ACPKIDEAL	FN	ST	S	VEL	TKH	M	849
Sc_Est2p	740	NNFHIRSK	SSK	GIF	RS	LIAL	FN	TR	772
Ea_p123	862	NINLR	IEG	I	LCT	LN	LM	Q	894

FIG. 42I

+



46/78

B.					
Sp_Tip1p	850	L	A	S	F
Sc_Est2p	773	L	N	S	T
Ea_p123	895	L	M	N	I
Sp_Tip1p	883	A	Q	A	Y
Sc_Est2p	794	-	-	-	Y
Ea_p123	928	Y	M	Q	A
Sp_Tip1p	916	L	A	E	I
Sc_Est2p	822	T	V	S	G
Ea_p123	961	T	R	A	F
Sp_Tip1p	949	F	K	Y	H
Sc_Est2p	855	T	S	-	-
Ea_p123	994	I	E	I	F
Sp_Tip1p	982	L	H	R	R
Sc_Est2p	878	I	Y	I	H
Ea_p123	1024	Q	S	L	I

FIG. 42J

+



47/78

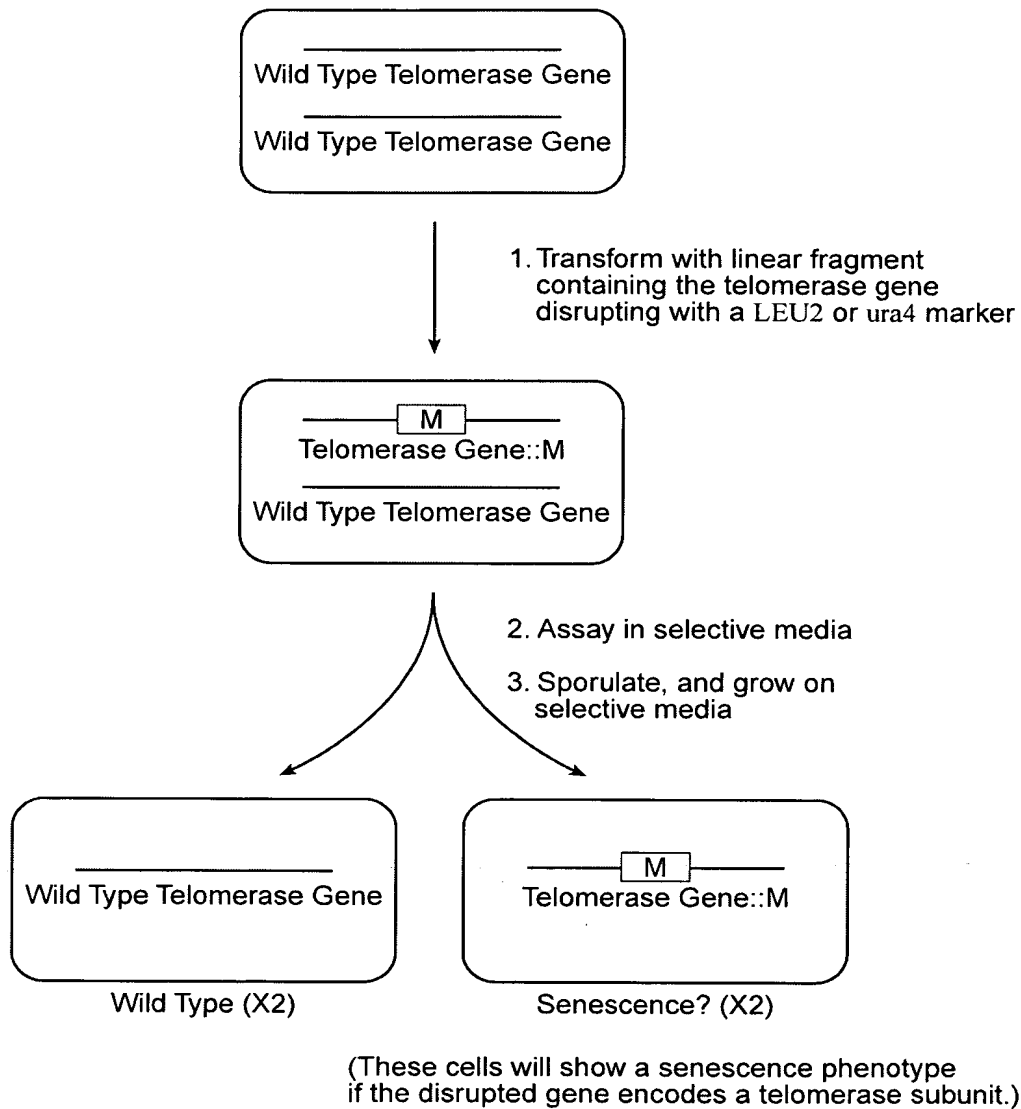
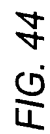


FIG. 43





49/78

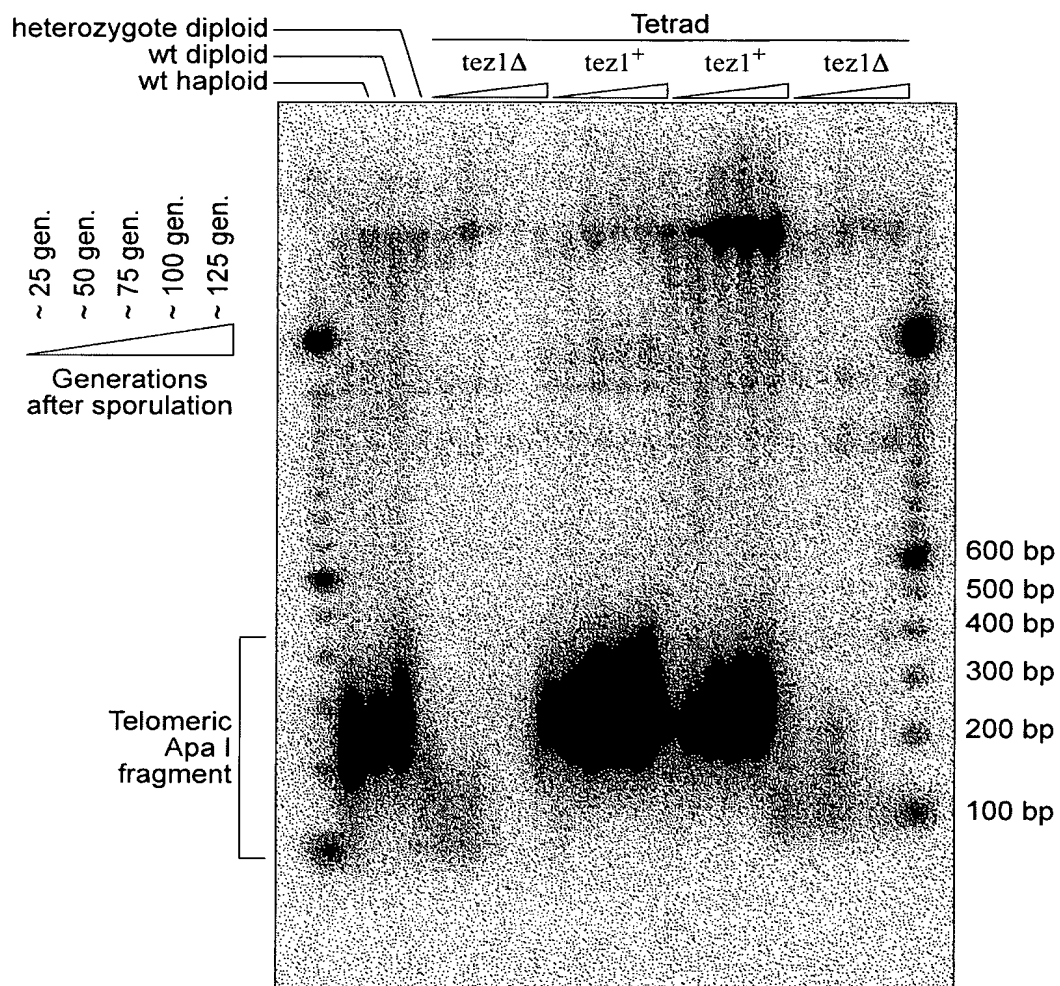


FIG. 45

+



50/78

```

1  ggtaccgatttacttcttcttccataagctaattgcttctcctcgaaacgctcctaaaatctctggaataatttttacaaga 80
81  actcaataacaataccaaagtaattccaatgaaggtgttaattagtagtgaatgaataattatttatttatttcggtgta 160
161  ccaagtataaggacaaaagaacaactcttcccccataaagacttttactttatttaatttacttttcaaatatatttcg 240
241  ggttcgcttacttttaactggtgactgttttagtctacttctagccaacccggtgtttctaccccgctcattggatat 320
321  agtcttggagtagctcacagaaatccctacaatcttctgatgagactatattagattcattacagtcctgcataattc 400
401  ttaacatggagccttacacttttagatgagtcacgcatgagagatatttggtatcatccaacggtttgccttgaaaaa 480
481  gttgataaattttgcaaaatcatgctccttagtgggttaacccggaagtttttggatgctgcacacgctctagcatg 560
561  attgagatatccaataatttctaccactacaactcctttaacggtttttattttctatttctcactcattgtgtt 640
641  ccaaatatgtatcatctcgtattaggctttttccggttttactcctggaaatcgtaacctttttcactattcccccaatga 720
721  ataatctaaattagtttcgcttaataattgataagtagaagagttggtgattcactcgtgtaattgtatttagtttaa 800
801  gatactttgcaaaacattttattagctatcattatataaaaaaacctataaattataaattatcaaatatttgoggtc 880
881  actattttataaacggttatgatcagtaggacacatttgcataatatagttatgcttaaatggttacttggtaactgc 958

959  ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
1 M T E H H T P K S R I L R F L E N Q Y V 20

1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
21 Y L C T L N D Y V Q L V L R G S P A S S 40

1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
41 Y S N I C E R L R S D V Q T S F S I F L 60

1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
61 H S T V V G F D S K P D E G V Q F S S P 80

1199 AAA TGC TCA CAG TCA GAG gtatatatattttgttttggatttttttctatttcgggtagtaatatatgggcag 1272
81 K C S Q S E 86

1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332
87 L I A N V V K Q M F D E S F E R R N L 106

1333 CTG ATG AAA GGG TTT TCC ATG gtaaggatttcttaattgtgaaatatttaccctgcaattactgtttcacaagaga 1405
107 L M K G F S M 113

1406 ttgtatttaacgcgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469
114 N H E D F R A M H V N G V Q N 128

```

FIG. 46A



51/78

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
 129 D L V S T F P N Y L I S I L E S K N W Q 148
 1530 CTT TTG TTA GAA AT gtaataaccggttaagatgttgcgcaactttgaacaagactgacaagtatag T ATC GGC 1601
 149 L L L E I I G 155
 1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
 156 S D A M H Y L L S K G S I F E A L P N D 175
 1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
 176 N Y L Q I S G I P L F K N N V F E T V 195
 1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
 196 S K K R K R T I E T S I T Q N K S A R K 215
 1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
 216 E V S W N S I S I S R F S I F Y R S S Y 235
 1842 AAG AAG TTT AAG CAA G gtaactaatactgttatcccttcataactaatttag AT CTA TAT TTT AAC 1907
 236 K K F K Q D L Y F N 245
 1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
 246 L H S I C D R N T V H M W L Q W I F P R 265
 1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
 266 Q F G L I N A F Q V K Q L H K V I P L V 285
 2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
 286 S Q S T V V P K R L L S K V Y N H Y C P Y I 305
 2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
 306 A K R L H R I S L S K V Y N H Y C P Y I 325
 2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
 326 D T H D D E K I L S Y S L K P N Q V F A 345
 2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
 346 F L R S I L V R V F P K L I I W G N Q R I 365

FIG. 46B



52/78

2268 TTT GAG ATA ATA TTA AAA G gatttgataaaattttattaccactaacgattttaccag AC CTC GAA ACT 2336
 366 F E I I L K D L E T 375
 2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F L K L S R Y E S F S L H Y L M S N I K 395
 2397 gtaatatgccaattttttttaccattaattaacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
 396 I S E I E W L V L G 405
 2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K R S N A K M C L S D F E K R K Q I F A 425
 2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445
 2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I T E S S D L R N R T V Y F R K D I W K 465
 2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L L C R P F I T S M K M E A F E K I N E 485
 2706 gtaatttaaagtatttttgcacaaagctaattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486 N N V R M D T Q K T 495
 2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
 496 T L P A V I R L L P K N T F R L I T 515
 2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtatttaatttttggatcatcaatgtactttacttctaattatta 2906
 516 N L R K R F L I K 524
 2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525 M G S N K K M L V S T N Q T L R P V 542
 2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A S I L K H L I N E S S G I P F N L E 562
 3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 46C



53/78

3089 tatataatgcgcgattcctcattattaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
 582 R K K Y F V R I D I 591
 3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
 592 K S C Y D R I K Q D L M F R I V K K L 611
 3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
 612 K D P E F V I R K Y A T I H A T S D R A 631
 3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttttttttcattggaattttttaacaa 3343
 632 T K N F V S E A F S Y F 643
 3344 attcttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
 644 D M V P F E K V V Q L L S M K T 659
 3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
 660 S D T L F V D F V D Y W T K S S E I F 679
 3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgtgaattgtataaca 3532
 680 K M L K E H L S G H I V K 692
 3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593
 693 I G N S Q Y L Q K V G I P Q G S 708
 3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
 709 I L S S F L C H F Y M E D L I D E Y L S 728
 3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
 729 F T K K K G S V L L R V V D D F L F I T 748
 3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc 3777
 749 V N K K D A K K F L N L S L R G 764
 3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
 765 F E K H N F S T S L E K T V 778
 3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900
 779 I N F E N S N G I I N N T F F N E S K K 798

FIG. 46D



54/78

```

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
799 R M P F F G F S V N M R S L D T L L A C 818

3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
819 P K I D E A L F N S T S V E L T K H M G 838

4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgttaactgaataatagctgacaataatcag A TCG 4089
839 K S F F Y K I L R S 848

4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
849 S L A S F A Q V F I D I T H N S K F N S 868

4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
869 C C N I Y R L G Y S M C M R A Q A Y L K 888

4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
889 R M K D I F I P Q R M F I T D 903

4275 aaagtcattaattaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
904 L L N V I G R K I W K K L A 917

4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
918 E I L G Y T S R R F L S S A E V K W 935

4402 ggctcgagacttcagcaaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
936 L F C L G M R D G L K 946

4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
947 P S F K Y H P C F E Q L I Y Q F Q S L T 966

4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA ATA 4588
967 D L I K P L R P V L R Q V L F L H R R I 986

4589 GCT GAT TAA tgtcattttcaattattattatatatacatcccttttattactgggtgtcttaacaataattattactaagtata 4665
987 A D * 989

```

FIG. 46E



55/78

4666 gctgacccccaaagcagcactataggatttctagtaaaagtaaaattaatctcgttattagttttgattgacttgct 4745
 4746 ttatcctttactttttaagaaagattgacagtgggtgctgactactgcccacatgccattaaacgggagtggttaaca 4825
 4826 ttaaaagtaatacatgaggttaattccttccatttagataaggaaagtgggtttctataatgaataatgcccgcacta 4905
 4906 atgcaaaaagacgaagattatcttctaaacaaaggggattaaagcataatcgaaaggaaagagagtaatatccccagtggt 4985
 4986 gttgaagaaagcaggataatttggaacagcttctgcagatgacaggctaaattttgtgaccgaattttggtaaaagc 5065
 5066 cccaggttatccatggcgccgcttgctactgagacgaaaaaactaaaggatagttgaataactaatagctcattta 5145
 5146 atgtcttataaaggtttgtttttcttgacttcaattttgcatgggtgaaaaaagaaatagtgtaagccattattggat 5225
 5226 tccgaaatagccaaatttttggttctcctcaaaagcggaaagtctaaagaacttatgaaagcttatgaggttcaaaaactcc 5305
 5306 tcctgatttaaaaggaggaatcttccacgatgaggaaatggatagcttatcagctgctgaggagaagcctaattttttgc 5385
 5386 aaaaaagaaatatcatgtggagacatctcttgatgaatcagatgaggagagatctccagcggatccttgatgtcaata 5465
 5466 acttctatttctgaaatgtatgggtcctactgtcgcttcgcttcgacttctcgtagctctacgcagttaagtgaacaaaggtacc 5544

FIG. 46F

$+$

+

FIG. 47A



57/78

210
gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220
gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

230
leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

240
arg asp gly leu leu leu arg leu val asp asp phe leu leu val
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

250
thr pro his leu thr his ala lys thr phe leu arg thr leu val
ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

260
arg gly val pro glu tyr gly cys val val asn leu arg lys thr
CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

270
val val asn phe pro val glu asp glu ala leu gly gly thr ala
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

280
phe val gln met pro ala his gly leu phe pro trp cys gly leu
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

290
leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

300
tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

310
phe lys ala gly arg asn met arg arg lys leu phe gly val leu
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

320
arg leu lys cys his ser leu phe leu asp leu gln val asn ser
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

330
leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

340
ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

350
410

FIG. 47B



58/78

420
gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430
arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

440
450
val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460
arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

470
480
thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490
pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

500
510
pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520
his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

530
540
thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550
arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

560
564
OP
TGA GTGAGTGT TTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC
CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTTCACTTCCCCAC
AGGCTGGCGTTCCGGTCCACCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACT
CCCCACATAGGAATAGTCCATCCCCAGATTGCGCCATTGTTACCCCTTCGCCCTGCCTTCC
TTTGCCTTCCACCCCCACCATTTCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTTTGGG
AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGG
GTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAATACTGAATATATGAGTT
TTTCAGTTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

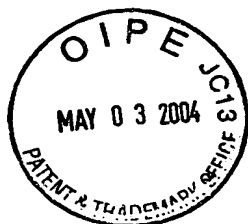
FIG. 47C



59/78

Motif -1	
Ep p123	...LVVSLIRCFFYVTEQQKSYSKT...
Sp Tez1	...FIIPILQSFFYITESSDLRNRT...
Sc Est2	...LIPKIIQTFFYCTEISSVTIV...
Hs TCP1	...YVVELLRSFFYVTETTFQKNRL...
consensus	FFY TE
Motif 0	
Ep p123	...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV... p hhh K hR h K R
Sp Tez1	...QKTTLPFAVIRLLPKKN--TFRITNLNRKRF... R PK R I
Sc Est2	...TSLNFNHSMRIIPKKSNNEFRIIAIPCRGAD...
Hs TCP1	...ARPALLTSRLRFIPKPD--GLRPVNMMDYVVG...
consensus	R PK R I
Motif A	
Ep p123	...PKLFFATMDIEKCYDSVNREKLSTFLK... h hDh GY h AF
Sp Tez1	...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
Sc Est2	...PELYFMKFDVKSCYDSIPRMECMRILK...
Hs TCP1	...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
consensus	F D YD
Motif B	
Ep p123	...NGKFYKQTKGIPQGLCVSSILSSFYYA... hPQG pS hh
Sp Tez1	...GNSQYLQKVGIPQGSILSSFLCHFYME...
Sc Est2	...EDKCYIREDGLFQGSSLSAPIVDLVYD...
Hs TCP1	...RATSYVQCQGIPQGSILSTLLCSLCYG...
consensus	G QG S
Motif C	
Ep p123	...PNVNLLMRLTDDYLLITTOENN... h F DD hhh Y
Sp Tez1	...KKGSVLLRVVDDFLFITVNKKD...
Sc Est2	...SQDTLILKLADDFLIISTDQQQ...
Hs TCP1	...RRDGLLLRLVDDFLLVTPHLTH...
consensus	DD L
Motif D	
Ep p123	...NVSRENGFKFNMKKL... Gh h cK
Sp Tez1	...LNLSLRGFEKHNFT...
Sc Est2	...KKLAMGGFQKYNKA...
Hs TCP1	...LRTLVRGVPEYGCVV...
consensus	G

FIG. 48



61/78

1 GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC
51 CCGCGATGCC GCGCGCTCCC CGCTGCCGAG CCGTGCGCTC CCTGCTGCGC
101 AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTTCGTGC GCGCCTGGG
151 GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG GGACCCGGCG GCTTTCGCG
201 CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC
251 CCCGCCGCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC
301 CCGAGTGCTG CANANGCTGT GCGANCGCG CGGAANAAC GTGCTGGCCT
351 TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCGCC CGAGGCCTTC
401 ACCACCAGCG TCGCAGCTA CCTGCCCAAC ACGGTGACCG ACGCACTGCG
451 GGGGAGCGGG GCGTGGGGGC TGCTGCTGCG CCGCGTGGGC GACGACGTGC
501 TGGTTACCT GCTGGCACGC TCGCGCNTNT TTGTGCTGGT GGNTCCCAGC
551 TCGCCTACC ANGTGTGCGG GCCGCCGCTG TACCAGCTCG GCGCTGCNAC
601 TCAGGCCCGG CCCCCGCCAC ACGCTANTGG ACCCGAANGC GTCTGGGATC
651 CAACGGGCCT GGAACCATAG GCTCAGGGAG GCCGGGGTCC CCCTGGGCTG
701 CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC
751 GTTGCCCAAG AGGCCAGGC GTGGCGCTGC CCCTGAGCCG GAGCGGACGC
801 CCGTTGGGCA GGGGTCTTGG GCCACCCGG GCAGGACGCC TGGACCGAGT
851 GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCCG AAGAAGCCAC
901 CTCTTTGGAG GTGCGCTCT CTGGCACGCG CCACTCCCAC CCATCCGTGG
951 GCCGCCAGCA CCACGCGGGC CCCCATCCA CATCGCGGCC ACCACGTCTC
1001 GGGACACGCC TTGTCCCCCG GTGTACGCCG AGACCAAGCA CTTCTCTAC
1051 TCCTCAGGCG ACAAGNACAC TGCGNCCCTC CTTCCTACTC AATATATCTG
1101 AGGCCAGCC TGA CTGGCGT TCGGGAGGTT CGTGAGACA NTCTTTCTGG
1151 TTCCAGGCCT TGGATGCCAG GATTCCCCCG AGGTTGCCCC GCCTGCCCCA
1201 GCGNTACTGG CAAATGCGGC CCTGTCTTCT GGAGCTGCTT GCGAACCACG
1251 CGCAGTGCCC CTACGGGGTG TTCTCAAGA CGCACTGCCC GCTGCGAGCT
1301 GCGGTACCC CAGCAGCCGG TGTCTGTGCC CGGGAGAAGC CCCAGGGCTC
1351 TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCCGTCG CCTGGTGCAG
1401 CTGCTCCGCC AGCACAGCAG CCCCTGGCAG GTGTACGGCT TCGTGCGGGC
1451 CTGCCCTGCG CGGCTGGTGC CCCAGGCCT CTGGGGCTCC AGGCACAACG
1501 AACGCCGCTT CTCAGGAAC ACCAAGAAGT TCATCTCCCT GCGGAAGCAT
1551 GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TCGGGGACTG
1601 CGCTTGGCTG CGCAGGAGCC CAGGGGTTGG CTGTGTTCCG GCCGCAGAGC
1651 ACCGTCTGCG TGAGGAGATC CTGGCCAAGT TCCTGCACTG GCTGATGAGT
1701 GTGTACGTCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCAC
1751 GTTTCAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTC TGGAGCAAGT
1801 TGCAAAGCAT TGGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGCGGGAG
1851 CTGTCCGAAG CAGAGGTCAG GCAGCATCGG GAAGCCAGGC CCGCCCTGCT
1901 GACGTCCAGA CTCCGCTTCA TCCCCAAGCC TGACGGGCTG CCGCCGATTG
1951 TGAACATGGA CTACGTCGTG GGAGCCAGAA CGTTCCGCAG AGAAAAGAGG
2001 GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTTCAGCG TGCTCAACTA
2051 CGAGCGGGCG CGGCGCCCCG GCCTCCTGGG CGCCTCTGTG CTGGGCCTGG
2101 ACGATATCCA CAGGGCCTGG CGCACCTTCG TGCTGCGTGT GCGGGCCAG
2151 GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA
2201 CGACACCATC CCCCAGGACA GGCTCACGGA GGTCACTGCC AGCATCATCA
2251 AACCCAGAA CACGTACTGC GTGCGTCGGT ATGCCGTGGT CCAGAAGGCC
2301 GCCCATTGGG ACGTCCGCAA GGCCTTCAAG AGCCACGTCT CTACCTTGAC
2351 AGACCTCCAG CCGTACATGC GACAGTTCGT GGCTCACCTG CAGGANAACA
2401 GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG
2451 GCCAGCAGTG GCCTCTTCGA CGTCTTCCTA CGCTTCATGT GCCACCACGC

FIG. 50A



62/78

2501 CGTGCGCATC AGGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG
2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
2601 AACAAAGCTGT TTGCGGGGAT TCGGCGGGAC GGGCTGCTCC TGC GTTTGGT
2651 GGATGATTTT TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
2701 TCAGGACCCCT GGTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACTTG
2751 CGGAAGACAG TGGTGAACCT CCCTGTAGAA GACGAGGCCC TGGGTGGCAC
2801 GGCTTTTGTT CAGATGCCGG CCCACGGCCT ATTCCCCTGG TCGGCGCTGC
2851 TGCTGGATAC CCGGACCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG
2951 GAGGAACATG CGTCGCAAAC TCTTTGGGGT CTGCGGCTG AAGTGTCA
3001 GCCTGTTTCT GGATTTGCAG GTGAACAGCC TCCAGACGGT GTGCACCAAC
3051 ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTCACG CATGTGTGCT
3101 GCAGCTCCCA TTTCATCAGC AAGTTTGGA GAACCCCA TTTTTCCTGC
3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
3201 AACGCAGGGA TGTCGCTGGG GGCCAAGGGC GCCGCCGCC CTCTGCCCTC
3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCTTGCTC AAGCTGACTC
3301 GACACCGTGT CACCTACGTG CCACTCCTGG GGTCACTCAG GACAGCCAG
3351 ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC
3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
3451 GGCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC
3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC CCAGGCCCGC
3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
3601 CCGCTGAAG GCTGAGTGTC CGGCTGAGGC CTGAGCGAGT GTCCAGCCAA
3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG
3701 CTCGGCTCCA CCCAGGGCC AGCTTTTCCT CACCAGGAGC CCGGCTTCCA
3751 CTCCCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC
3801 GCCCTGCCCT CCTTTGCCTT CCACCCAC CATCCAGGTG GAGACCTGA
3851 GAAGGACCCCT GGGAGCTCTG GGAATTTGGA GTGACCAAAG GTGTGCCCTG
3901 TACACAGGCG AGGACCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
3951 GGGGGGAGGT GCTGTGGGAG TAAAATACTG AATATATGAG TTTTTCAGTT
4001 TTGAAAAAAA AAAAAAAAAA AAAAAAAAAA

FIG. 50B



63/78

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCGATGCC
1 -----+-----+-----+-----+-----+ 60
CGTCGCGACGCAGGACGACGCGTGCACCCTTCGGGACCGGGCGGTTGGGGGCGCTACGG

a A A L R P A A H V G S P G P G H P R D A -
b Q R C V L L R T W E A L A P A T P A M P -
c S A A S C C A R G K P W P R P P P R C R -

GCGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCT
61 -----+-----+-----+-----+-----+ 120
CGCGCGAGGGGCGACGGCTCGGCACGCGAGGACGACGCGTCGGTGATGGCGTCCACGA

a A R S P L P S R A L P A A Q P L P R G A -
b R A P R C R A V R S L L R S H Y R E V L -
c A L P A A E P C A P C C A A T T A R C C -

GCCGCTGGCCACGTTTCGTGCGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGACGCGCG
121 -----+-----+-----+-----+-----+ 180
CGGCGACCGGTGCAAGCACGCGCGGACCCCGGGTCCCGACCGCCGACCACGTCGCGCC

a A A G H V R A A P G A P G L A A G A A R -
b P L A T F V R R L G P Q G W R L V Q R G -
c R W P R S C G A W G P R A G G W C S A G -

GGACCCGGCGGCTTTCCGCGCGNTGGTGGCCANTGCGTGGTGCGTGCCTTGGGANGN
181 -----+-----+-----+-----+-----+ 240
CCTGGGCGCGGAAAGGCGCGNACCACCGGGTNACGNACCACGACGCGGACCCCTNCN

a G P G G F P R ? G G P ? ? G V R A L G ? -
b D P A A F R A ? V A ? C ? V C V P W ? ? -
c T R R L S A R W W P ? A W C A C P G ? ? -

ANGGNGCCCCCGCGCCCCCTCCTTCGCGCAGGTGTCCTGCCTGAANGANCTGGTGGC
241 -----+-----+-----+-----+-----+ 300
TNCCGNCGGGGGGCGGCGGGGAGGAAGGCGGTCCACAGGACGGACTTNCCTNGACCACCG

a ? A A P R R P L L P P G V L P E ? ? G G -
b ? ? P P A A P S F R Q V S C L ? ? L V A -
c G ? P P P P P P S A R C P A * ? ? W W P -

CCGAGTGCTGCANANGCTGTGCGANCGCGGCGCAANAACGTGCTGGCCTTCGGCTTCGC
301 -----+-----+-----+-----+-----+ 360
GGCTCACGACGTNTNCGACACGCTNGCGCCGCGCTTNTTGCACGACCGGAAGCCGAAGCG

a P S A A ? A V R ? R R E ? R A G L R L R -
b R V L ? ? L C ? R G A ? N V L A F G F A -
c E C C ? ? C A ? A A R ? T C W P S A S R -

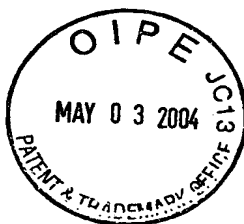
GCTGCTGGACGGGGCCCGCGGGGGCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTA
361 -----+-----+-----+-----+-----+ 420
CGACGACCTGCCCGGGCGCCCCGGGGGGGCTCCGGAAGTGGTGGTTCGACGCGTCGAT

a A A G R G P R G P P R G L H H Q R A Q L -
b L L D G A R G G P P E A F T T S V R S Y -
c C W T G P A G A P P R P S P P A C A A T -

CCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGCTGCG
421 -----+-----+-----+-----+-----+ 480
GGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTCGCCCCGACCCCCGACGACGACGC

a P A Q H G D R R T A G E R G V G A A A A -
b L P N T V T D A L R G S G A W G L L L R -
c C P T R * P T H C G G A G R G G C C C A -

FIG. 51A



64/78

a P R G R R R A G S P A G T L R ? ? C A G -
b R V G D D V L V H L L A R C A ? F V L V -
c A W A T T C W F T C W H A A R ? L C W W -

GGNTCCCGCTGCGCTACCGTGTGCGGGCCGCGCTGTACCAGCTCGGCGCTGCNAC
541 -----+-----+-----+-----+-----+ 600
CCNAGGGTCGACGCGGATGGTNCACACGCCCGGCGGACATGGTCGAGCCGCGACGNTG

a G S Q L R L P ? V R A A A V P A R R C ? -
b ? P S C A Y ? V C G P P L Y Q L G A A T -
c ? P A A P T ? C A G R R C T S S A L ? L -

TCAGGCCCCGGCCCCGCCACACGCTANTGGACCCGAANGCGTCTGGGATCCAACGGGCCCT
601 -----+-----+-----+-----+-----+ 660
AGTCCGGGCGGGGGCGGTGTGCGATNACCTGGGCTTNCGCAGACCCTAGGTTGCCCGGA

a S G P A P A T R ? W T R ? R L G S N G P -
b Q A R P P P H A ? G P E ? V W D P T G L -
c R P G P R H T L ? D P ? A S G I Q R A W -

GGAACCATAGCGTACGGGAGGCGGGGTCCCCCTGGGCTGCCAGCCCCGGGTGCGAGGAG
661 -----+-----+-----+-----+-----+ 720
CCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCCGACGGTCGGGGCCCCACGCTCCTC

a G T I A S G R P G S P W A A S P G C E E -
b E P * R Q G G R G P P G L P A P G A R R -
c N H S V R E A G V P L G C Q P R V R G G -

GCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTGTCCCAAGAGGCCAGGCGTGCGCTGC
721 -----+-----+-----+-----+-----+ 780
CGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGGGTTCTCCGGGTCCGCACCGCGACG

a A R G Q C Q P K S A V A Q E A Q A W R C -
b R G G S A S R S L P L P K R P R R G A A -
c A G A V P A E V C R C P R G P G V A L P -

CCCTGAGCCGAGCGGACGCCCCGTTGGGCAGGGGTCTGGGCCACCCGGGCAGGACGCC
781 -----+-----+-----+-----+-----+ 840
GGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAGGACCCGGGTGGGCCCCGTCTGCGG

a P * A G A D A R W A G V L G P P G Q D A -
b P E P E R T P V G Q G S W A H P G R T P -
c L S R S G R P L G R G P G P T R A G R L -

TGGACCGAGTGACCGTGGTTTCTGTGTGGTGTACCTGCCAGACCCGCCGAAGAAGCCAC
841 -----+-----+-----+-----+-----+ 900
ACCTGGCTCACTGGCACCAAGACACACCACAGTGGACGGTCTGGGCGGCTTCTTCGGTG

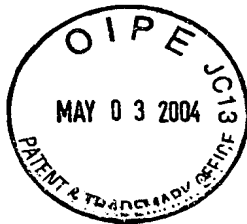
a W T E * P W F L C G V T C Q T R R R S H -
b G P S D R G F C V V S P A R P A E E A T -
c D R V T V V S V W C H L P D P P K K P P -

CTCTTTGGAGGTGCGCTCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCGGCCAGCA
901 -----+-----+-----+-----+-----+ 960
GAGAAACCTCCCACGCGAGAGACCGTGC GCGGTGAGGGTGGGTAGGCACCCGGCGGTCTGT

a L F G G C A L W H A P L P P I R G P P A -
b S L E G A L S G T R H S H P S V G R Q H -
c L W R V R S L A R A T P T H P W A A S T -

CCACGCGGGCCCCCATCCACATCGCGGCCACCACGTCTCTGGGACACGCCTTGTCCCCCG
961 -----+-----+-----+-----+-----+ 1020
GGTGCGCCCGGGGGTAGGTGTAGCGCCGTTGGTGCAGGACCTGTGCGGAACAGGGGGC

FIG. 51B



65/78

a P R G P P I H I A A T T S W D T P C P P -
b H A G P P S T S R P P R P G T R L V P R -
c T R A P H P H R G H H V L G H A L S P G -

1021 GTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACAAGNACACTGCGNCCCTC
-----+-----+-----+-----+ 1080
CACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAGTCCGCTGTTTCNTGTGACGCNGGGAG

a V Y A E T K H F L Y S S G D K ? T A ? L -
b C T P R P S T S S T P Q A T ? T L R P S -
c V R R D Q A L P L L L R R Q ? H C ? P P -

1081 CTTCCTACTCAATATATCTGAGGCCCGAGCTGACTGGCGTTCGGGAGGTTTCGTGGAGACA
-----+-----+-----+-----+ 1140
GAAGGATGAGTTATATAGACTCCGGGTTCGACTGACCGCAAGCCCTCCAAGCACCTCTGT

a L P T Q Y I * G P A * L A F G R F V E T -
b F L L N I S E A Q P D W R S G G S W R ? -
c S Y S I Y L R P S L T G V R E V R G D ? -

1141 NTCTTTCTGGTTCAGGCCCTTGGATGCCAGGATTCCTCCGCGAGGTTGCCCCGCGCTGCCCCA
-----+-----+-----+-----+ 1200
NAGAAAGACCAAGGTCCGGAACCTACGGTCTTAAGGGGCGTCCAACGGGGCGGACGGGGT

a ? F L V P G L G C Q D S P Q V A P P A P -
b S F W F Q A L D A R I P R R L P R L P Q -
c L S G S R P W M P G F P A G C P A C P S -

1201 GCGNTACTGGCAAATGCGGCCCCCTGTTTCTGGAGCTGCTTGGGAACACGCGCAGTGCCCC
-----+-----+-----+-----+ 1260
CGCNATGACCGTTTACGCCGGGGACAAAGACCTCGACGAACCTTGGTGCGCGTCACGGG

a A ? L A N A A P V S G A A W E P R A V P -
b R Y W Q M R P L F L E L L G N H A Q C P -
c ? T G K C G P C F W S C L G T T R S A P -

1261 CTACGGGGTGTTCCTCAAGACGCACTGCCCCGTGCGAGCTGCGGTACCCCCAGCAGCCGG
-----+-----+-----+-----+ 1320
GATGCCCCACAAGGAGTTCTGCGTGACGGGCGACGCTCGACGCCAGTGGGGTTCGTGGGCC

a L R G V P Q D A L P A A S C G H P S S R -
b Y G V F L K T H C P L R A A V T P A A G -
c T G C S S R R T A R C E L R S P Q Q P V -

1321 TGTCTGTGCCCCGGGAGAAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGAACACAG
-----+-----+-----+-----+ 1380
ACAGACACGGGCCCTCTTCGGGGTCCCGAGACACCGCCGGGGGCTCCTCCTCCTTGTGTC

a C L C P G E A P G L C G G P R G G G T Q -
b V C A R E K P Q G S V A A P E E E E H R -
c S V P G R S P R A L W R P P R R R N T D -

1381 ACCCCCGTCGCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCT
-----+-----+-----+-----+ 1440
TGGGGGCAGCGGACCACGTGACGAGGCGGTGCTGTGTCGCGGGACCGTCCACATGCCGA

a T P V A W C S C S A S T A A P G R C T A -
b P P S P G A A A P P A Q Q P L A G V R L -
c P R R L V Q L L R Q H S S P W Q V Y G F -

1441 TCGTGCGGGCCTGCCTGCGCCGGCTGGTGGCCCCAGGCCTCTGGGGCTCCAGGCACAACG
-----+-----+-----+-----+ 1500
AGCACGCCCCGACGACGCGGCCGACCACGGGGGTCCGGAGACCCCGAGGTCCGTGTTGC

FIG. 51C



66/78

a S C G P A C A G W C P Q A S G A P G T T -
b R A G L P A P A G A P R P L G L Q A Q R -
c V R A C L R R L V P P G L W G S R H N E -

1501 AACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCT
-----+-----+-----+-----+ 1560
TTGCGGCGAAGGAGTCCTTGTGGTTCTTCAAGTAGAGGGACCCCTTCGTACGGTTCGAGA

a N A A S S G T P R S S S P W G S M P S S -
b T P L P Q E H Q E V H L P G E A C Q A L -
c R R F L R N T K K F I S L G K H A K L S -

1561 CGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGCTTGGCTGCGCAGGAGCC
-----+-----+-----+-----+ 1620
GCGACGTCCTCGACTGCACCTTCTACTCGCACGCCCTGACGCGAACCGACGCGTCTCGG

a R C R S * R G R * A C G T A L G C A G A -
b A A G A D V E D E R A G L R L A A Q E P -
c L Q E L T W K M S V R D C A W L R R S P -

1621 CAGGGGTTGGCTGTGTTCCGGCCGCAGAGCACCGTCTGCGTGAGGAGATCCTGGCCAAGT
-----+-----+-----+-----+ 1680
GTCCCCAACCGACACAAGGCCGGCGTCTCGTGGCAGACGCACTCTCTAGGACCGGTTCA

a Q G L A V F R P Q S T V C V R R S W P S -
b R G W L C S G R R A P S A * G D P G Q V -
c G V G C V P A A E H R L R E E I L A K F -

1681 TCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCA
-----+-----+-----+-----+ 1740
AGGACGTGACCGACTACTCACACATGCACAGCTCGACGAGTCCAGAAAGAAAATACAGT

a S C T G * * V C T S S S C S G L S F M S -
b P A L A D E C V R R R A A Q V F L L C H -
c L H W L M S V Y V V E L L R S F F Y V T -

1741 CGGAGACCACGTTTCAAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGT
-----+-----+-----+-----+ 1800
GCCTCTGGTGCAAAGTTTCTTGTCCGAGAAAAAGATGGCCTTCTCACAGACCTCGTTCA

a R R P R F K R T G S F S T G R V S G A S -
b G D H V S K E Q A L F L P E E C L E Q V -
c E T T F Q K N R L F F Y R K S V W S K L -

1801 TGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTCCGAAG
-----+-----+-----+-----+ 1860
ACGTTTCGTAACCTTAGTCTGTCTGTAACCTTCTCCACGTCGACGCCCTCGACAGCCTTC

a C K A L E S D S T * R G C S C G S C R K -
b A K H W N Q T A L E E G A A A G A V G S -
c Q S I G I R Q H L K R V Q L R E L S E A -

1861 CAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCA
-----+-----+-----+-----+ 1920
GTCTCCAGTCCGTCGTAGCCCTTCGGTCCGGGCGGGACGACTGCAGGTCTGAGCGGAAGT

a Q R S G S I G K P G P P C * R P D S A S -
b R G Q A A S G S G A R P A D V Q T P L H -
c E V R Q H R E A R P A L L T S R L R F I -

1921 TCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTGCTGGGAGCCAGAA
-----+-----+-----+-----+ 1980
AGGGGTTTCGACTGCCCCAGCCCGGCTAACACTTGTACCTGATGCAGCACCCCTCGGTCTT

FIG. 51D



67/78

```

a   S P S L T G C G R L * T W T T S W E P E -
b   P Q A * R A A A D C E H G L R R G S Q N -
c   P K P D G L R P I V N M D Y V V G A R T -

CGTTCGCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTTCAGCG
1981 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040
GCAAGGCGTCTCTTTCTCCCGGCTCGCAGAGTGGAGCTCCCACTTCCGTGACAAGTCGC

a   R S A E K R G P S V S P R G * R H C S A -
b   V P Q R K E G R A S H L E G E G T V Q R -
c   F R R E K R A E R L T S R V K A L F S V -

TGCTCAACTACGAGCGGGCGCGGCCCGGCCTCCTGGGCGCCTCTGTGCTGGGCCTGG
2041 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100
ACGAGTTGATGCTCGCCCGCGCCGCGGGCCGAGGACCCGCGGAGACACGACCCGGACC

a   C S T T S G R G A P A S W A P L C W A W -
b   A Q L R A G A A P R P P G R L C A G P G -
c   L N Y E R A R R P G L L G A S V L G L D -

ACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCAGGACCCGCCGC
2101 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2160
TGCTATAGGTGTCCCGGACCGCGTGAAGCACGACGACACGCCCCGGGTCTGGGCGGGC

a   T I S T G P G A P S C C V C G P R T R R -
b   R Y P Q G L A H L R A A C A G P G P A A -
c   D I H R A W R T F V L R V R A Q D P P P -

CTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCAGGACA
2161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2220
GACTCGACATGAAACAGTTCCACCTACACTGCCCGCGCATGCTGTGGTAGGGGTCTGT

a   L S C T L S R W M * R A R T T P S P R T -
b   * A V L C Q G G C D G R V R H H P P G Q -
c   E L Y F V K V D V T G A Y D T I P Q D R -

GGCTCACGGAGGTCATCGCCAGCATCATCAAACCCAGAACACGTACTGCGTGCGTCCGT
2221 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2280
CCGAGTGCCTCCAGTAGCGGTCTAGTAGTTTGGGGTCTTGTGCATGACGCACGCAGCCA

a   G S R R S S P A S S N P R T R T A C V G -
b   A H G G H R Q H H Q T P E H V L R A S V -
c   L T E V I A S I I K P Q N T Y C V R R Y -

ATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAGCCACGTCT
2281 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2340
TACGGCACCAGGTCTTCCGGCGGGTACCCGTGCAGGCGTTCCGGAAGTTCTCGGTGCAGA

a   M P W S R R P P M G T S A R P S R A T S -
b   C R G P E G R P W A R P Q G L Q E P R L -
c   A V V Q K A A H G H V R K A F K S H V S -

CTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGANAACA
2341 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400
GATGGAAGTGTCTGGAGGTCCGCATGTACGCTGTCAAGCACCGAGTGGACGTCTNTTGT

a   L P * Q T S S R T C D S S W L T C R ? T -
b   Y L D R P P A V H A T V R G S P A G ? Q -
c   T L T D L Q P Y M R Q F V A H L Q ? N S -

GCCCCGTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGCAGTG
2401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2460
CGGGCGACTCCCTACGGCAGCAGTAGCTCGTCTCGAGGAGGGACTTACTCCGGTCTGCAC

```

FIG. 51E



68/78

a A R * G M P S S S S R A P P * M R P A V -
b P A E G C R R H R A E L L P E * G Q Q W -
c P L R D A V V I E Q S S S L N E A S S G -

GCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGT
2461 -----+-----+-----+-----+-----+ 2520
CGGAGAAGCTGCAGAAGGATGCGAAGTACACGGTGGTGC GGACGCGTAGTCCCCGTTCA

a A S S T S S Y A S C A T T P C A S G A S -
b P L R R L P T L H V P P R R A H Q G Q V -
c L F D V F L R F M C H H A V R I R G K S -

CCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCC
2521 -----+-----+-----+-----+-----+ 2580
GGATGCAGGTACGGTCCCCTAGGGCGTCCCGAGGTAGGAGAGGTGCGACGAGACGTCCG

a P T S S A R G S R R A P S S P R C S A A -
b L R P V P G D P A G L H P L H A A L Q P -
c Y V Q C Q G I P Q G S I L S T L L C S L -

TGTGCTACGGCGACATGGAGAACAAGCTGTTTGC GGGGATTCCGGCGGACGGGTGCTCC
2581 -----+-----+-----+-----+-----+ 2640
ACACGATGCCGCTGTACCTCTTGTTCGACAAACGCCCTAAGCCGCCCTGCCCGACGAGG

a C A T A T W R T S C L R G F G G T G C S -
b V L R R H G E Q A V C G D S A G R A A P -
c C Y G D M E N K L F A G I R R D G L L L -

TGCGTTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCACGCGAAAACCTTCC
2641 -----+-----+-----+-----+-----+ 2700
ACGCAAACCACCTACTAAAGAACAACCACTGTGGAGTGGAGTGGGTGCGCTTTTGAAGG

a C V W W M I S C W * H L T S P T R K P S -
b A F G G * F L V G D T S P H P R E N L P -
c R L V D D F L L V T P H L T H A K T F L -

TCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACCTTGC GGAAGACAG
2701 -----+-----+-----+-----+-----+ 2760
AGTCCTGGGACCAGGCTCCACAGGGA CTATACCGACGCACCACTTGAACGCCTTCTGTCT

a S G P W S E V S L S M A A W * T C G R Q -
b Q D P G P R C P * V W L R G E L A E D S -
c R T L V R G V P E Y G C V V N L R K T V -

TGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGATGCCGG
2761 -----+-----+-----+-----+-----+ 2820
ACCACTTGAAGGGACATCTTCTGCTCCGGGACCCACCGTGCCGAAAACAAGTCTACGGCC

a W * T S L * K T R P W V A R L L F R C R -
b G E L P C R R R G P G W H G F C S D A G -
c V N F P V E D E A L G G T A F V Q M P A -

CCCACGGCCTATTCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGA
2821 -----+-----+-----+-----+-----+ 2880
GGGTGCCGGATAAGGGGACCACGCCGACGACGACCTATGGGCCTGGGACCTCCACGTCT

a P T A Y S P G A A C C W I P G P W R C R -
b P R P I P L V R P A A G Y P D P G G A E -
c H G L F P W C G L L L D T R T L E V Q S -

GCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCT
2881 -----+-----+-----+-----+-----+ 2940
CGCTGATGAGGTGATACGGGCCTGGAGGTAGTCTCGGTCAGAGTGGAAGTTGGCGCCGA

FIG. 51F



69/78

a A T T P A M P G P P S E P V S P S T A A -
b R L L Q L C P D L H Q S Q S H L Q P R L -
c D Y S S Y A R T S I R A S L T F N R G F -

2941 TCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGTGTCACA
-----+-----+-----+-----+-----+ 3000
AGTTCCGACCCTCCTTGTACGCAGCGTTTGAGAAACCCAGAACGCCGACTTCACAGTGT

a S R L G G T C V A N S L G S C G * S V T -
b Q G W E E H A S Q T L W G L A A E V S Q -
c K A G R N M R R K L F G V L R L K C H S -

3001 GCCTGTTTCTGGATTTCAGGTTGAACAGCCTCCAGACGGTGTGCACCAACATCTACAAGA
-----+-----+-----+-----+-----+ 3060
CGGACAAAGACCTAAACGTCCACTTGTGCGAGGTCTGCCACACGTGGTTGTAGATGTTCT

a A C F W I C R * T A S R R C A P T S T R -
b P V S G F A G E Q P P D G V H Q H L Q D -
c L F L D L Q V N S L Q T V C T N I Y K I -

3061 TCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTCATCAGC
-----+-----+-----+-----+-----+ 3120
AGGAGGACGACGTCCGCATGTCCAAAGTGCCTACACACGACGTGCGAGGGTAAAGTAGTCG

a S S C C R R T G F T H V C C S S H F I S -
b P P A A G V Q V S R M C A A A P I S S A -
c L L L Q A Y R F H A C V L Q L P F H Q -

3121 AAGTTTGAAGAACCCACATTTTCTGCGCGTCATCTCTGACACGGCCTCCCTCTGCT
-----+-----+-----+-----+-----+ 3180
TTCAAACCTTCTTGGGGTGTA AAAAGGACGCGCAGTAGAGACTGTGCCGGAGGGAGACGA

a K F G R T P H F S C A S S L T R P P S A -
b S L E E P H I F P A R H L * H G L P L L -
c V W K N P T F F L R V I S D T A S L C Y -

3181 ACTCCATCCTGAAAGCCAAGAACGCAGGGATGTGCTGCGGGCCAAGGGCGCCGCCGCC
-----+-----+-----+-----+-----+ 3240
TGAGGTAGGACTTTCGGTTCTTGCGTCCCTACAGCGACCCCGGTTCCCGCGGCGGCCG

a T P S * K P R T Q G C R W G P R A P P A -
b L H P E S Q E R R D V A G G Q G R R P -
c S I L K A K N A G M S L G A K G A A G P -

3241 CTCTGCCCTCCGAGGCCGTGTCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGCTGACTC
-----+-----+-----+-----+-----+ 3300
GAGACGGGAGGCTCCGGCACGTACCGACACGGTGGTTTCGTAAGGACGAGTTCGACTGAG

a L C P P R P C S G C A T K H S C S S * L -
b S A L R G R A V A V P P S I P A Q A D S -
c L P S E A V Q W L C H Q A F L L K L T R -

3301 GACACCGTGTACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGCAGCTGA
-----+-----+-----+-----+-----+ 3360
CTGTGGCACAGTGGATGCACGGTGAGGACCCAGTGAGTCTGTGCGGTCTGCGTCCGACT

a D T V S P T C H S W G H S G Q P R R S * -
b T P C H L R A T P G V T Q D S P D A A E -
c H R V T Y V P L L G S L R T A Q T Q L S -

3361 GTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGAGCCAACCCGGCACTGC
-----+-----+-----+-----+-----+ 3420
CAGCCTTTCGAGGGCCCTGCTGCGACTGACGGGACCTCCGCGTGGTTGGGCCGTGACG

FIG. 51G



70/78

a V G S S R G R R * L P W R P Q P T R H C -
b S E A P G D D A D C P G G R S Q P G T A -
c R K L P G T T L T A L E A A A N P A L P -

3421 CCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCACAGCCAGGCCGAGAGCA
-----+-----+-----+-----+-----+ 3480
GGAGTCTGAAGTTCTGGTAGGACCTGACTACCGGTGGGCGGGTGTCTGGTCCGGTCTCTGT

a P Q T S R P S W T D G H P P T A R P R A -
b L R L Q D H P G L M A T R P Q P G R E Q -
c S D F K T I L D * W P P A H S Q A E S R -

3481 GACACCAGCAGCCCTGTACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCACAC
-----+-----+-----+-----+-----+ 3540
CTGTGGTCTGTCTGGGACAGTGGCGGCCGAGATGCAGGGTCCCTCCCTCCCCCGCGGGTGTG

a D T S S P V T P G S T S Q G G R G G P H -
b T P A A L S R R A L R P R E G G A A H T -
c H Q Q P C H A G L Y V P G R E G R P T P -

3541 CCAGGCCCCGACCGCTGGGAGTCTGAGGCTGAGTGAGTGTGTTGGCCGAGGCCATGATGT
-----+-----+-----+-----+-----+ 3600
GGTCCGGGCGTGGCGACCCTCAGACTCCGACTCACTCACAAACCGGCTCCGGACGTACA

a P G P H R W E S E A * V S V W P R P A C -
b Q A R T A G S L R P E * V F G R G L H V -
c R P A P L G V * G L S E C L A E A C M S -

3601 CCGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGCGAGTGTCCAGCCAAGGGCTGAGTG
-----+-----+-----+-----+-----+ 3660
GGCCGACTTCCGACTCACAGGCCGACTCCGACTCGCTCACAGGTCTGGTTCCCGACTCAC

a P A E G * V S G * G L S E C P A K G * V -
b R L K A E C P A E A * A S V Q P R A E C -
c G * R L S V R L R P E R V S S Q G L S V -

3661 TCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCAGGGCC
-----+-----+-----+-----+-----+ 3720
AGGTCTGTGTGGACGGCAGAAGTGAAGGGTGTCCGACCGGAGCCGAGGTGGGGTCCCGG

a S S T P A V F T S P Q A G A R L H P R A -
b P A H L P S S L P H R L A L G S T P G P -
c Q H T C R L H F P T G W R S A P P Q G Q -

3721 AGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGA
-----+-----+-----+-----+-----+ 3780
TCGAAAAGGAGTGGTCTCGGGCCGAAGGTGAGGGGTGTATCCTTATCAGGTAGGGGTCT

a S F S S P G A R L P L P T * E * S I P R -
b A F P H Q E P G F H S P H R N S P S P D -
c L F L T R S P A S T P H I G I V H P Q I -

3781 TTCGCCATTGTTACCCCTCGCCCTGCCCTCCTTTGCCCTTCCACCCCAACCATCCAGGTG
-----+-----+-----+-----+-----+ 3840
AAGCGGTAACAAGTGGGAGCGGGACGGGAGGAAACGGAAGGTGGGGTGGTAGGTCCAC

a F A I V H P S P C P P L P S T P T I Q V -
b S P L F T P R P A L L C L P P P P S R W -
c R H C S P L A L P S F A F H P H H P G G -

3841 GAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTG
-----+-----+-----+-----+-----+ 3900
CTCTGGGACTCTTCTCGGACCCTCGAGACCCTTAAACCTCACTGGTTTCCACACGGGAC

FIG. 51H



71/78

a E T L R R T L G A L G I W S D Q R C A L -
b R P * E G P W E L W E F G V T K G V P C -
c D P E K D P G S S G N L E * P K V C P V -

3901 TACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGT
-----+-----+-----+-----+-----+ 3960
ATGTGTCCGCTCCTGGGACGTGGACCTACCCCCAGGGACACCCAGTTTAACCCCCCTCCA

a Y T G E D P A P G W G S L W V K L G G G -
b T Q A R T L H L D G G P C G S N W G E V -
c H R R G P C T W M G V P V G Q I G G R C -

3961 GCTGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTGAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+ 4020
CGACACCCTCATTTTATGACTTATATACTCAAAAAGTCAAACTTTTTTTTTTTTTTTTTT

a A V G V K Y * I Y E F F S F E K K K K K -
b L W E * N T E Y M S F S V L K K K K K K -
c C G S K I L N I * V F Q F * K K K K K K -

4021 AAAAAAAAAA
----- 4029
TTTTTTTTT

a K K K -
b K K -
c K K -

FIG. 51I

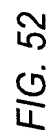


FIG. 52



73/78

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG

10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

30
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

40
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

50
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

60
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

70
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

80
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

90
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

100
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

110
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

120
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

130
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

FIG. 53A



74/78

200 210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

230 240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

260 270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

280
val val ser pro ala arg pro ala glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

290 300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

320 330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350 360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

370
glu thr ile phe leu gly ser arg pro pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

380 390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

410 420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

FIG. 53B



75/78

430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CCG

470
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

480
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

490
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

500
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

510
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

520
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

530
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

540
arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

550
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

560
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

570
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CCG

580
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

FIG. 53C



76/78

650
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

660
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

670
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

680
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

690
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

700
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

710
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

720
his gly his val arg lys ala phe lys ser his val ser thr leu
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG

730
thr asp leu gln pro tyr met arg gln phe val ala his leu gln
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG

740
glu thr ser pro leu arg asp ala val val ile glu gln ser ser
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC

750
ser leu asn glu ala ser ser gly leu phe asp val phe leu arg
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC

760
phe met cys his his ala val arg ile arg gly lys ser tyr val
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC

770
gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC

780
cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG

790
ile arg arg asp gly leu leu leu arg leu val asp asp phe leu
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG

800
810
820
830
840
850
860
870

FIG. 53D



77/78

880
 leu val thr pro his leu thr his ala lys thr phe leu arg thr
 TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

890
 leu val arg gly val pro glu tyr gly cys val val asn leu arg
 CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG

910
 lys thr val val asn phe pro val glu asp glu ala leu gly gly
 AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

920
 thr ala phe val gln met pro ala his gly leu phe pro trp cys
 ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC

940
 gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
 GGC CTG CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

950
 ser ser tyr ala arg thr ser ile arg ala ser leu thr phe asn
 TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC

960
 arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
 CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

970
 val leu arg leu lys cys his ser leu phe leu asp leu gln val
 GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

980
 asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
 AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

990
 leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
 CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

1000
 his gln gln val trp lys asn pro thr phe phe leu arg val ile
 CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

1010
 ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
 TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

1020
 ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
 GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GCC GGC CCT CTG CCC

1030
 ser glu ala val gln trp leu cys his gln ala phe leu leu lys
 TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

1040
 leu thr arg his arg val thr tyr val pro leu leu gly ser leu
 CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

1050
 1060
 1070
 1080
 1090

FIG. 53E



78/78

1100 1110
 arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
 AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

1120
 leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
 CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG GCA CTG CCC TCA GAC

1130 1132
 phe lys thr ile leu asp OP
 TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCGAGAGCAGA

CACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCACACCC
 AGGCCCCGACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCC
 GGCTGAAGGCTGAGTGTCGGGCTGAGGCCTGAGCGAGTGTCAGCCAAGGGCTGAGTGTC
 CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAG
 CTTTTCYTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT
 CGCCATTGTTACCCYTCGCCCTGCCYTCCTTTGCCTTCCACCCCCACCATCCAGGTGGA
 GACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTA
 CACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGC
 TGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTTGRAAAAAAAAAAAAAAAAAA
 AAAAAAAAAA

FIG. 53F

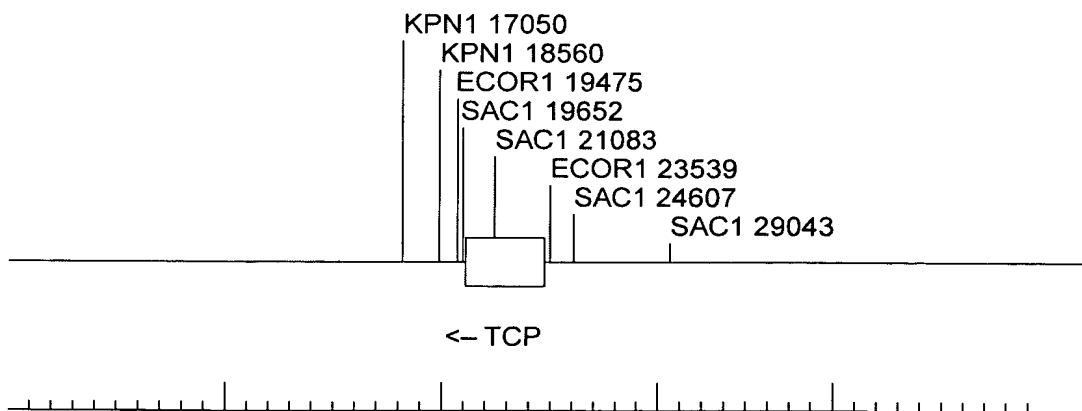


FIG. 54